

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:30:13 ; Search time 71 Seconds

(without alignments)
106,976 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTKDFNLDTVSXKDSGA.....ALMCKNKATATCHCSIHVSK 57

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298	98.0	57	11	AA05238
2	298	98.0	57	17	AA05267
3	298	98.0	57	20	AA05665
4	298	98.0	57	20	AA05308
5	291	95.7	57	13	AA028298
6	291	95.7	57	14	AA041280
7	291	95.7	57	14	AA033850
8	291	95.7	57	14	AA033912
9	291	95.7	57	20	AA106570
10	185	60.9	34	15	AA062635

11	185	60.9	34	19	AA066443
12	185	60.9	34	21	AA01742
13	185	60.9	34	23	AA030978
14	181	59.5	41	18	AA031659
15	178	58.6	34	14	AA043070
16	176	57.9	34	20	AA084352
17	157	51.6	34	11	AA05237
18	157	51.6	34	19	AA066445
19	157	51.6	34	20	AA03209
20	157	51.6	34	21	AA01744
21	128	42.1	39	18	AA031658
22	105	34.5	34	14	AA050953
23	104	34.2	34	14	AA037314
24	104	34.2	34	14	AA043272
25	104	34.2	34	15	AA066162
26	104	34.2	34	17	AA05263
27	97	31.9	34	13	AA028299
28	97	31.9	34	17	AA05245
29	97	31.9	34	23	AA05775
30	97	31.9	34	23	AA05777
31	96	31.6	34	14	AA050954
32	96	31.6	34	23	AA05776
33	94.5	31.1	52	10	AA098498
34	94.5	31.1	52	14	AA037315
35	94.5	31.1	52	20	AA043430
36	94.5	31.1	52	20	AA03210
37	93	30.6	34	21	AA011032
38	92	27.0	34	14	AA039311
39	82	27.0	34	14	AA041281
40	82	27.0	34	14	AA030171
41	82	27.0	47	16	AA085080
42	81.5	26.8	35	18	AA031660
43	72	23.7	46	23	AB029366
44	63	20.7	33	22	AB027639
45	61.5	20.2	63	22	AB020484

ALIGNMENTS

RESULT 1
AA05238
ID AA05238 standard; protein: 57 AA.
AC AA05238:
XX 04-AUG-1990 (first entry)
XX
XX Nisin precursor protein and leader peptide encoded by DNA derived from
DE Streptococcus lactis ATCC 11454.
XX
XX Nisin precursor peptide; nisin leader peptide;
KW post-translational modification; Streptococcus lactis ATCC 11454;
XX
XX Streptococcus lactis ATCC 11454.
FH Key
FT Peptide
FT /note="leader fragment responsible for inducing post-
FT translational modification"
FT Protein
FT /note="nisin precursor peptide"
XX
XX WO900558-A.
XX
XX 25-JUN-1990.
XX
XX 30-JUN-1989; 89MO-US02820.
XX
XX 05-JUL-1988; 88US-0214959.
XX
XX (UYMA-) UNIV OF MARYLAND.

Cationic peptide n
Cationic peptide N
Transplant media a
Subtilin-nisin chl
Lactococcus lactis
Peptide QSP-91241
Subtilin precursor
Cationic peptide s
Amino acid sequenc
Nisin-subtilin chl
Nisin A/SSA. Lact
Epidermin. Staphy
Nisin A. Lactococ
Sequence of the ba
Nisin A. Lactococ
Sequence of nisin
Nisin Z. Lactococ
Lactococcus lactis
Lactococcus lactis
Nisin A/H270. Lac
Lactococcus lactis
Sequence of pre-ep
Epla protein. Sta
S. epidermis readi
Amino acid sequenc
Lactococcus lactis
Lactococcus bacter
Bacteriocin (Gene
Bacteriocin LL-2.
Hyacin M51 (lanthl
Mutant lantibiotic
Streptococcus poly
Novel human diagn
Novel human diagn

PI Hangen NJ;
 XX WPI; 1990-051685/07.
 DR N-PSDB; AAO93354.
 XX
 PT Leader peptide sequence -
 PT including post-translational modification of polypeptide(s)
 XX
 XX Disclosure; Fig 3; 19pp; English.
 XX
 CC The leader peptide assists in inducing post-translational modification in
 CC a protein precursor when attached to the precursor as a leader. The
 CC precursor polypeptide contains Ser, Thr and Cys which undergo
 CC modification after translation to arrive at the mature protein,
 CC having unusual amino acids. Tag a is claimed in the patent.
 XX
 SQ Sequence 57 AA;
 Query Match 98.0%; Score 298; DB 11; Length 57;
 Best Local Similarity 98.2%; Pred. No. 8.7e-29;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
 DB
 RESULT 2
 AAR95267
 ID AAR95267 standard; Protein; 57 AA.
 XX
 AC AAR95267;
 XX
 DT 01-AUG-1996 (first entry)
 XX
 DE Pre-nisin A.
 XX
 KM Nisin A; nisa gene; antimicrobial; preservative; antibiotic;
 KM lantibiotic; protein engineering.
 XX
 OS Lactococcus lactis strain NIZO R5.
 XX
 XX
 FH Key
 FT 1..23 Location/Qualifiers
 FT Peptide /label= sig.peptide
 XX
 PN W09616180-A1.
 XX
 PN 30-MAY-1996.
 XX
 XX 20-NOV-1995; 95WO-GB02699.
 XX
 PR 19-NOV-1994; 94GB-0023404.
 XX
 PA (BIOT-) BIOTECHNOLOGY & BIOLOGICAL SCI RES COUNC.
 XX
 PI Dodd HM, Gasson MJ;
 XX
 DR WPI; 1996-268616/27.
 DR N-PSDB; AAT9660.
 XX
 XX Making cell which expresses nisin but does not contain natural nisa
 PT gene - by providing cell with variant nisa gene, and genes for nisin
 PT modification, secretion and immunity
 PT
 XX Disclosure; Fig 7; 69pp; English.
 XX
 XX The gene cluster nisABTCTPRK (see AAT9660 and AAT9661) of Lactococcus
 CC lactis includes the nisa gene coding for pre-nisin A (AAR95267, see
 CC also AAR95263) and the genes for nisin modification, secretion and
 CC immunity. nisa (AAR95268) and nisc (AAR95270) are believed to be
 CC involved in reactions that modify pre-nisin; nist (AAR95269) is
 CC similar to a transport ATPase and is involved in translocation of

CC nisin out of the cell; nist (AAR95271) is involved in immunity to
 CC nisin. Replacement of the natural, chromosomal copy of the nisa
 CC gene with a variant nisa gene allows produ. of high levels of nisin
 CC A variants in Lactococcus lactis hosts.
 XX
 SQ Sequence 57 AA;
 Query Match 98.0%; Score 298; DB 17; Length 57;
 Best Local Similarity 98.2%; Pred. No. 8.7e-29;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
 DB
 RESULT 3
 AAY06665
 ID AAY06665 standard; Protein; 57 AA.
 XX
 AC AAY06665;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE Nisin A of Lactobacillus lactis.
 XX
 KM Nisin A; nisa gene; animal performance; growth rate;
 KM feed conversion; probiotic; bacteriocin.
 XX
 OS Lactobacillus lactis.
 XX
 PN W09941978-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 12-FEB-1999; 99WO-IB00250.
 XX
 PR 18-FEB-1998; 98GB-0003424.
 XX
 PA (PFIZ) PFIZER INC.
 PA (PFIZ) PFIZER LTD.
 XX
 PI Flanagan AJ, Haxell MA, Rolph TP;
 XX
 DR WPI; 1999-527402/44.
 DR N-PSDB; AAX87792.
 XX
 PT Novel performance enhancing method resulting in improved growth
 PT rates and feed conversion efficiencies in animals
 PT
 XX Disclosure; Page 61; 79pp; English.
 XX
 XX The present sequence represents the nisin A precursor peptide of
 CC Lactobacillus lactis NIZO R5. The sequence is deduced from the
 CC nisa A gene (see AAX87792). Nisin A is lantibiotic-containing
 CC bacteriocin. The invention relates to methods of enhancing
 CC performance in an animal by administering a bacterium capable of
 CC expressing a performance enhancing polypeptide such as nisin A or
 CC nisin Z. The enhanced performance results in improved growth rates
 CC and feed conversion efficiencies. The bacterium, which may be
 CC genetically modified to express the performance enhancing
 CC polypeptide, is administered to the gastrointestinal tract,
 CC especially to an embryo or neonatal animal. If the polypeptide is
 CC nisin, it may also inhibit ruminal methane, decrease acetate to
 CC propionate ratios and prevent amino acid deamination.
 CC Administration of an appropriate bacterium which may act as a
 CC probiotic may also help control enteric pathogens in poultry. The
 CC polypeptide is produced continuously in the gut, maintaining a
 CC constant level. The protein is eventually degraded, leaving no
 CC residues in the meat.
 XX
 SQ Sequence 57 AA;

Query Match 98.0%; Score 298; DB 20; Length 57;
 Best Local Similarity 98.2%; Pred. No. 8.7e-29;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MSTKDFNLDLVSVSKKDSGASPRITSTISICTPGCKTGALMGCMNMTATCHCSIHSK 57
 1 MSTKDFNLDLVSVSKKDSGASPRITSTISICTPGCKTGALMGCMNMTATCHCSIHSK 57

RESULT 4

AA03208
 ID AA03208 standard; Protein: 57 AA.

AC AA03208;

DT 03-AUG-1999 (first entry)

DE Amino acid sequence of nisin A.

Sublancin 168; antimicrobial; food preservative; pro-sublancin 168;
 Gram-positive bacteria; pre-sublancin 168; nisin A.

OS unknown.

PN WO9903352-A1.

PD 28-JAN-1999.

PF 17-JUL-1998; 98WO-US14547.

PR 18-JUL-1997; 97US-0053035.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Hansen JN;

DR WPI: 1999-131752/11.

PT New antimicrobial peptide, sublancin 168, from *Bacillus subtilis* -
 used for, e.g. treatment of infections caused by Gram negative
 bacteria and as food preservative

PS Disclosure: Page 53; 71pp; English.

CC This is the amino acid sequence of nisin A used in the method of
 the invention involving the use of prosublancin 168. The peptide
 designated sublancin 168, is an antimicrobial useful for treating
 infections and preserving food against spoilage bacteria,
 particularly Gram-positive bacteria. Pro-sublancin 168 and
 pre-sublancin 168, are the precursors of sublancin 168. Sublancin
 168 is very stable at low pH and can be autoclaved without damage.
 It does not decompose after 2 years in aqueous solution of about
 neutral pH.

CC Sequence 57 AA;

Query Match 98.0%; Score 298; DB 20; Length 57;
 Best Local Similarity 98.2%; Pred. No. 8.7e-29;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MSTKDFNLDLVSVSKKDSGASPRITSTISICTPGCKTGALMGCMNMTATCHCSIHSK 57
 1 MSTKDFNLDLVSVSKKDSGASPRITSTISICTPGCKTGALMGCMNMTATCHCSIHSK 57

RESULT 5

AA028298
 ID AA028298 standard; Protein: 57 AA.

AC AA028298;

DT 02-APR-1993 (first entry)

XX

DE Sequence encoded by nisz gene isolated from *L. lactis* NIZO 22186.
 XX Lantibiotic; nisin Z; nisin A; analogue; food preservative.
 XX Lactococcus lactis.

OS Lactococcus lactis.

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT /label= leader

PN WO9218633-A.

PD 29-OCT-1992.

PF 09-APR-1992; 92WO-NL00068.

PR 11-APR-1991; 91NL-0000634.

PA (NIZO-) NIZO STICHTING NEDERLANDS INST ZUIVELOND.

PI De Vos WM, Kuipers OP, Slezzen RJ;

DR WPI: 1992-382116/46.

DN N-PSDB; AA028299.

PT New lantibiotic cpds. related to nisin A - and Lactococcus
 strains which produce them, useful as preservatives for foods and
 animal feeds

PS Example: Fig 3; 42pp; English.

CC Total DNA was isolated from the *L. lactis* strains NIZO 22186 and
 CC NIZO R5. With the aid of the nisa gene of strain NIZO R5 as a probe,
 CC a 4.5 kb HindIII fragment was identified in the total DNA strain
 CC NIZO 22186, which was then cloned in M13 mp18. The DNA sequence of
 CC the gene for nisin Z production (nisZ) was determined by making use
 CC of oligos complementary to the 5' and 3' flanking sequences of the
 CC nisa gene. The nucleotide sequence of the nisz gene is found to be
 CC identical to that of the nisa gene with the exception of a C to A
 CC transversion in posn. 148 which results in the replacement of a C to A
 CC His27 by Asn27. The above indicate that the structure of nisin Z is
 CC as shown in AA028299.

CC Sequence 57 AA;

Query Match 95.7%; Score 291; DB 13; Length 57;
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 1 MSTKDFNLDLVSVSKKDSGASPRITSTISICTPGCKTGALMGCMNMTATCHCSIHSK 57
 1 MSTKDFNLDLVSVSKKDSGASPRITSTISICTPGCKTGALMGCMNMTATCNCISIHSK 57

RESULT 6

AA041280
 ID AA041280 standard; Protein: 57 AA.

AC AA041280;

DT 10-MAR-1994 (first entry)

DE Bacteriocin.

Bacteriocin; LL-2; gram positive bacteria; antimicrobial; food;
 inhibit.

OS Lactococcus lactis.

FT Key Location/Qualifiers

FT Peptide 1..23
 FT /label= sig-peptide
 FT Protein 24..57

FT /label= mat_protein
 FT /note= "Claim 3"
 FT Modified-site
 FT /note= "The CH2 in the side chain joins
 FT the S in residue 30"
 FT Modified-site
 FT /note= "The S joins the side chain of residue 26"
 FT Modified-site
 FT /note= "The CH2 in the side chain joins
 FT the S in residue 34"
 FT Modified-site
 FT /note= "The S joins the side chain of residue 31"
 FT Modified-site
 FT /note= "The CH2 in the side chain joins
 FT the S in residue 42"
 FT Modified-site
 FT /note= "The S joins the side chain of residue 36"
 FT Modified-site
 FT /note= "The CH2 in the side chain joins
 FT the S in residue 51"
 FT Modified-site
 FT /note= "The S joins the side chain of residue 46"
 FT Modified-site
 FT /note= "The CH2 in the side chain joins
 FT the S in residue 52"
 FT Modified-site
 FT /note= "The S joins the side chain of residue 49"
 FT US5232849-A.
 FT 03-AUG-1993.
 FT 01-JUL-1991; 91US-0721774.
 FT 01-JUL-1991; 91US-0721774.
 FT 14-MAY-1992; 92US-0882079.
 FT (QUES-) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
 FT Henderson JT, Marugg JD, Van Wassenaar PD, Vedamuthu ER;
 FT WPI; 1993-287077/36.
 FT N-PSDB; AAQ49150.
 FT Bacteriocin from Lactococcus lactis subspecies lactis - useful as
 FT inhibitory against Gram-positive bacteria
 FT Disclosure: Page 13-14 (col 15,16,17,18); 14pp; English.
 FT The sequence (AAQ49150) was amplified using primers (AAQ49151-52). The
 FT encoded protein inhibits selected gram positive bacteria and this
 FT property is enhanced if further purified by HPLC. The materials
 FT being treated to provide inhibition are preferably foods, although
 FT other materials may be treated.
 FT Sequence 57 AA;
 SQ
 Query Match 95.7%; Score 291; DB 14; Length 57;
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
 DB 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
 RESULT 7
 AAR33850
 ID AAR33850 standard; Protein; 57 AA.
 XX
 AC AAR33850;
 XX
 PF 12-MAY-1993 (first entry)
 DT

XX Bacteriocin LL-2 precursor.
 DE
 XX Food treatment.
 KW
 RW Lactococcus lactis sub-species lactis NRRL B-18809.
 OS
 XX Lactococcus lactis (subspecies lactis).
 XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Peptide 24..57
 FT /note= "mature peptide"
 FT US5173297-A.
 FT 22-DEC-1992.
 FT 01-JUL-1991; 91US-0721774.
 FT 01-JUL-1991; 91US-0721774.
 FT (QUES-) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
 FT Henderson JT, Marugg JD, Van Wassenaar PD, Vedamuthu ER;
 FT WPI; 1993-017533/02.
 FT N-PSDB; AAQ34782.
 FT Inhibition of Gram-positive bacteria - using bacteriocin derived
 FT from Lactococcus lactis sub-species lactis NRRL B-18809
 FT Disclosure: Page 13; 14pp; English.
 FT The sequence is that of bacteriocin LL-2 precursor which can be
 FT used in a method for the inhibition of Gram-positive bacteria.
 FT LL-2 is especially useful for treatment of food, although other
 FT non-food materials may also be treated.
 FT Sequence 57 AA;
 SQ
 Query Match 95.7%; Score 291; DB 14; Length 57;
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
 DB 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCSTHYSK 57
 RESULT 8
 AAR39312
 ID AAR39312 standard; Protein; 57 AA.
 XX
 AC AAR39312;
 XX
 DT 21-JAN-1994 (first entry)
 XX Lactococcal bacteriocin polypeptide precursor.
 DE
 XX Bacteriocin; inhibition; polypeptide; Lactococcus lactis.
 KW
 XX Lactococcus lactis (subspecies lactis).
 XX Key Location/Qualifiers
 FT Peptide 24..57
 FT /label= Bacteriocin.
 FT US5231165-A.
 FT 27-JUL-1993.
 XX
 PD 01-JUL-1991; 91US-0721774.
 XX
 PF
 DT

PR 01-JUL-1991; 91US-0721774.
 PR 14-MAY-1992; 920S-0882715.
 XX
 PA (QUES-) QUEST INT FLAVORS & FOOD INGREDIENTS CO.
 XX
 PI Henderson JT, Marugg JD, Vanwassenaar PD, Vedamuthu ER;
 DR WPI: 1993-249768/31.
 DR N-PSDB; AA046618.
 XX
 PT Isolated and purified polypeptide from *Lactococcus lactis* sub
 PT species *lactis* - has inhibitory activity against gram-positive
 PI bacteria for e.g. food etc.
 XX
 PS Claim 1; Column 17-18; 13pp; English.
 XX
 CC The isolated bacteriocin obtained from the polypeptide precursor has
 CC an inhibitory activity against selected gram positive bacteria.
 CC The amount of bacteriocin required to provide inhibition is 15-100
 CC arbitrary units per gram of material. The materials being treated
 CC with the bacteriocin to provide inhibition are especially foodstuffs.
 SQ Sequence 57 AA:
 Query Match 95.7%; Score 291; DB 14; Length 57;
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIHVS 57
 DB 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIHVS 57
 RESULT 9
 ID AAY06670 standard; Protein: 57 AA.
 XX AAY06670;
 XX AC
 XX 09-NOV-1999 (first entry)
 DT
 XX
 DE Nisin Z of *Lactobacillus lactis*.
 XX
 KW Nisin Z; nls Z; lantibiotic; animal performance; growth rate;
 KW feed conversion; bacteriocin; probiotic.
 XX
 OS *Lactobacillus lactis*.
 XX
 MO9941978-A1.
 PD 26-AUG-1999.
 XX
 XX 12-FEB-1999; 99WO-IB00250.
 XX
 PR 18-FEB-1998; 98GB-0003424.
 XX
 PA (PRIZ) PRIZER INC.
 PA (PRIZ) PRIZER LTD.
 PI Flanagan AJ, Haxell MA, Rolph TP;
 DR WPI: 1999-527402/44.
 DR N-PSDB; AAX87793.
 XX
 FT Novel performance enhancing method resulting in improved growth
 PT rates and feed conversion efficiencies in animals
 XX
 PS Disclosure: Page 71; 79pp; English.
 XX
 CC This sequence represents the nisin Z peptide of *Lactobacillus*
 CC *lactis* strain 22186. Nisin Z is a natural analogue of nisin A (see
 CC AAY06665), a lanthionine-containing bacteriocin. The invention
 CC relates to methods of enhancing performance in an animal by

CC administering a bacterium capable of expressing a performance
 CC enhancing polypeptide such as nisin A or nisin Z. The enhanced
 CC performance results in improved growth rates and feed conversion
 CC efficiencies. The bacterium, which may be genetically modified to
 CC express the performance enhancing polypeptide, is administered to
 CC the gastrointestinal tract, especially to an embryo or neonatal
 CC animal. If the polypeptide is nisin, it may also inhibit ruminal
 CC methane, decrease acetate to propionate ratios and prevent amino
 CC acid denaturation. Administration of an appropriate bacterium which
 CC may act as a probiotic may also help control enteric pathogens in
 CC poultry. The polypeptide is produced continuously in the gut,
 CC maintaining a constant level. The protein is eventually degraded,
 CC leaving no residues in the meat.
 SQ Sequence 57 AA:
 Query Match 95.7%; Score 291; DB 20; Length 57;
 Best Local Similarity 96.5%; Pred. No. 6.2e-28;
 Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 OY 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIHVS 57
 DB 1 MSTKDFNDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCNCISIHVS 57
 RESULT 10
 ID AAR62635 standard; peptide; 34 AA.
 XX AAR62635;
 XX AC
 XX 15-JUN-1995 (first entry)
 DT
 XX
 DE Putative intermediate for lanthionine-contg. peptide.
 XX
 KW Lanthionine; methylanthionine; lantibiotic; antiviral;
 KW immunosuppressant; antimicrobial; enzyme inhibitor.
 XX
 OS Synthetic.
 XX
 JP06253885-A.
 PD 13-SEP-1994.
 XX
 PF 09-MAR-1993; 93JP-0048385.
 XX
 PR 09-MAR-1993; 93JP-0048385.
 XX
 PA (AJIN) AJINOMOTO KK.
 PA (AJIN) AJINOMOTO KK.
 DR WPI: 1994-329026/41.
 XX
 XX Prepn. of lanthionine contg. peptide(s) - useful as antimicrobial,
 PT antiviral drugs, immunosuppressants and enzyme inhibitors
 XX
 PS Example 2; Page 7; 8pp; Japanese.
 XX
 CC This is one of 5 peptides (AAR62635-R63639) containing Cys and Ser or
 CC Thr residues which were synthesised and tested for their usefulness
 CC as intermediates for the preparation of peptides which include
 CC lanthionine. Peptides 3 and 4 (AAR62637-8) produced lanthionine, while
 CC both lanthionine and methylanthionine could be produced from peptide
 CC 2 (AAR62636).
 CC
 SQ Sequence 34 AA:
 Query Match 60.9%; Score 185; DB 15; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2.4e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 24 ITSTSLCTPGCKTGALMGCMKMTATCNCISIHVS 57
 DB 1 ITSTSLCTPGCKTGALMGCMKMTATCNCISIHVS 34

RESULT 11
 AAM66443
 ID AAM66443 standard; peptide; 34 AA.
 XX
 AC AAM66443;
 XX
 DT 12-JUN-1999 (first entry)
 XX
 DE Cationic peptide nisin.
 XX
 KM Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KM bacterial infection; tolerance; antibacterial; microorganism;
 KM bacteria; fungus; parasite; virus.
 XX
 OS Lactococcus lactis.
 XX
 PN WO9840401-A2.
 XX
 DT 17-SEP-1998.
 XX
 DR 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, Mc Nicol PJ, West MHP;
 DR WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Disclosure; Page 10; 105pp; English.
 XX
 CC AAM66393 to AAM66469 represent native cationic peptides from the
 CC present invention. The present invention describes compositions and
 CC methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 XX
 SQ Sequence 34 AA;
 XX
 Query Match 60.9%; Score 185; DB 19; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2,4e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 24 ITSTSICPGCKTGALMGCMNKTATCHCSIHVSK 57
 ||| ||||||||||||||||||||||||||||
 DB 1 ITSTSICPGCKTGALMGCMNKTATCHCSIHVSK 34
 XX
 RESULT 12
 AAY91742
 ID AAY91742 standard; Peptide; 34 AA.
 XX
 AC AAY91742;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Cationic peptide Nisin amino acid sequence.

XX
 KM Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KM leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KM multidrug resistance.
 XX
 OS Unidentified.
 XX
 PN WO9965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 XX 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;
 XX
 DR WPI; 2000-223549/19.
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours
 XX
 PS Disclosure; Page 11; 94pp; English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 XX
 SQ Sequence 34 AA;
 XX
 Query Match 60.9%; Score 185; DB 21; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2,4e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 24 ITSTSICPGCKTGALMGCMNKTATCHCSIHVSK 57
 ||| ||||||||||||||||||||||||||||
 DB 1 ITSTSICPGCKTGALMGCMNKTATCHCSIHVSK 34
 XX
 RESULT 13
 AAU90978
 ID AAU90978 standard; Peptide; 34 AA.
 XX
 AC AAU90978;
 XX
 DT 05-JUN-2002 (first entry)
 XX
 DE Transplant media associated antimicrobial peptide #14.
 XX
 KM Transplant; antimicrobial peptide; pore forming agent;
 KM cell surface receptor binding compound; kidney transplant;
 KM cardioplegia; organ transplant; transplant rejection.
 XX
 OS Lactococcus lactis.
 XX
 PN WO200209738-A1.
 XX
 PD 07-FEB-2002.
 XX
 XX 27-JUL-2001; 2001WO-US23785.
 XX
 PF 28-JUL-2000; 2000US-221632P.
 PR 17-NOV-2000; 2000US-249602P.

PR 15-MAY-2001; 2001US-290932P.
 XX
 PA (MURP/) MURPHY C J.
 XX
 PI Murphy CJ, Reid TW, Meanulty JF;
 XX
 DR WPI; 2002-268995/31.
 XX
 PT Media comprising antimicrobial polypeptides or pore forming agents
 PT and/or cell surface receptor binding compounds useful for the storage
 PT and preservation of organs prior to transplant
 XX
 PS Disclosure; Page 25; 78pp; English.
 XX
 CC The invention describes new transplant compositions comprising
 CC antimicrobial polypeptides or pore forming agents and/or cell surface
 CC receptor binding compounds. The media is capable of extending the
 CC preservation period past 72 hours and can provide organs with increased
 CC functionality upon transplant. Animals receiving kidneys stored in the
 CC media of the present invention for either three or four days had serum
 CC creatinine levels of less than half of those observed in control animals
 CC receiving kidneys stored in UW solution (defined in the specification)
 CC alone. Lower serum creatinine levels are indicative of healthier kidneys
 CC and a more preferable prognosis for the transplant patient. The media of
 CC the invention are useful for decreasing the incidence and/or severity of
 CC delayed graft function in patients receiving transplanted kidneys stored
 CC and/or treated in the media. The media may also be used in procedures
 CC such as cardioplegia. It is contemplated that transplant of healthier
 CC organs leads to a decrease in chronic rejection. This sequence represents
 CC an antimicrobial peptide studied in the development of the transplant
 CC media.
 XX
 SQ Sequence 34 AA;
 OY Query Match 60.9%; Score 185; DB 23; Length 34;
 Best Local Similarity 97.1%; Pred. No. 2.4e-15;
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 DB 24 ITSTSLCTPGCKTGALMGCMNMTATCHCSIHVSK 57
 1 ITSTSLCTPGCKTGALMGCMNMTATCHCSIHVSK 34
 RESULT 14
 AAY31659
 ID AAY31659 standard; Protein; 41 AA.
 XX
 AC AAY31659;
 DE 09-NOV-1999 (first entry)
 XX
 DE Subtilin-nisin chimera.
 XX
 KW Nisin; subtilin; lantibiotic; chimera; mutant; bacteriocide;
 KM preservative.
 XX
 OS Chimeric - Lactococcus lactis.
 OS Chimeric - Bacillus subtilis.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..7
 FT /note= "signal peptide"
 FT Protein 8..41
 FT /note= "mature protein"
 FT Region 8..18
 FT /note= "nisin (1-11)"
 FT Region 19..41
 FT /note= "subtilin(12-32)"
 XX
 MO9711713-A1.
 PD 03-APR-1997.
 XX

PF 30-SEP-1996; 96MO-US15160.
 XX
 PR 28-SEP-1995; 95US-0535494.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Hansen JN;
 XX
 DR WPI; 1997-225847/20.
 DR N-PSDB; AAX87829.
 XX
 PT Lantibiotic mutants and chimera(s) - having enhanced stability and
 PT activity compared to nisin
 XX
 PS Example; Fig 2; 60pp; English.
 XX
 CC The present sequence represents a chimeric pre-peptide composed of
 CC a subtilin leader region and a subtilin-nisin fusion comprising
 CC residues 1-11 of Bacillus subtilis subtilin and residues 12-32 of
 CC Lactococcus lactis nisin. The subtilin-nisin fusion was not
 CC processed into a functional lantibiotic when expressed in B.
 CC subtilis. A heterogeneous mixture of products was produced, with
 CC none of the products having the expected properties of a correctly
 CC processed polypeptide. However, the mixture contained a minor
 CC component with a specific activity that exceeded that of nisin.
 CC The invention provides lantibiotic mutants and chimeras (see also
 CC AAY31658) having enhanced activity and stability compared to nisin
 CC and subtilin. They can be produced by cultivation of transformed
 CC host cells and used e.g. as food preservatives to treat, kill or
 CC inhibit the growth of microorganisms and/or their spores.
 XX
 SQ Sequence 41 AA;
 OY Query Match 59.5%; Score 181; DB 18; Length 41;
 Best Local Similarity 80.5%; Pred. No. 9.1e-15;
 Matches 33; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 DB 17 DSGASPRITSTSLCTPGCKTGALMGCMNMTATCHCSIHVSK 57
 1 DSKITPQMKSSSPCTPGCKTGALMGCMNMTATCHCSIHVSK 41
 RESULT 15
 AAR43070
 ID AAR43070 standard; peptide; 34 AA.
 XX
 AC AAR43070;
 DE 09-JUN-1994 (first entry)
 XX
 DE Lactococcus lactis bacteriocin.
 XX
 KW Lactococcus; bacteriocin; lactobacillus casei; food; antimicrobial;
 KW inhibition; contamination.
 XX
 OS Lactococcus lactis.
 OS
 XX
 PN EP573768-A.
 PD 15-DEC-1993.
 XX
 PF 28-APR-1993; 93EP-0106911.
 PR 08-MAY-1992; 92US-0880003.
 XX
 PA (UNIL) QUEST INT BV.
 XX
 PI Henderson JT, Vandenberg PA, Vedamuthu ER;
 XX
 DR WPI; 1993-396506/50.
 DR
 XX
 PT New Lactococcus sp. for preserving foods - contains DNA encoding
 PT bacteriocin from donor L-lactis strain and recipient L lactis

PT	Strain
1	0.00
2	0.00
3	0.00
4	0.00
5	0.00
6	0.00
7	0.00
8	0.00
9	0.00
10	0.00
11	0.00
12	0.00
13	0.00
14	0.00
15	0.00
16	0.00
17	0.00
18	0.00
19	0.00
20	0.00
21	0.00
22	0.00
23	0.00
24	0.00
25	0.00
26	0.00
27	0.00
28	0.00
29	0.00
30	0.00
31	0.00
32	0.00
33	0.00
34	0.00
35	0.00
36	0.00
37	0.00
38	0.00
39	0.00
40	0.00
41	0.00
42	0.00
43	0.00
44	0.00
45	0.00
46	0.00
47	0.00
48	0.00
49	0.00
50	0.00
51	0.00
52	0.00
53	0.00
54	0.00
55	0.00
56	0.00
57	0.00
58	0.00
59	0.00
60	0.00
61	0.00
62	0.00
63	0.00
64	0.00
65	0.00
66	0.00
67	0.00
68	0.00
69	0.00
70	0.00
71	0.00
72	0.00
73	0.00
74	0.00
75	0.00
76	0.00
77	0.00
78	0.00
79	0.00
80	0.00
81	0.00
82	0.00
83	0.00
84	0.00
85	0.00
86	0.00
87	0.00
88	0.00
89	0.00
90	0.00
91	0.00
92	0.00
93	0.00
94	0.00
95	0.00
96	0.00
97	0.00
98	0.00
99	0.00
100	0.00

PS Claim 14; Page 14; 15pp; English.

CC The sequence is of a bacteriocin isolated from *L. lactis*. The
CC peptide can be used to preserve food products against the growth of
CC *L. bulgaricus*, *L. casei* and other contaminants. The bacteriocin is
CC nisin-like in its method of inhibition.

Sequence . 34 AA;

Query Match

Query Match	58.6%;	Score 178;	DB 14;	Length 34;
Best Local Similarity	94.1%;	Pred. No. 1.7e-14;		
Matches 32; Conservative	1;	Mismatches 1;	Indels	

Best Local Similarity 94.1%; Pred. No. 1.7e-14;
Matches 32; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
OY      24 ITSTSCTPGCKTGALMGCMKKTATCHCSIHYSK   57  
          ||| | | | | | | | | | | : | | | |  
Db       1 ITSISLCTPGCKTGTALMGCMKTFATCNCISIHYSK   34
```

Db 1 ITSISLCTPGCKTGALMGCMKTATCNCSIHVS 34

```

Search completed: June 7, 2003, 15:31:45
Load time : 75 secs

```

Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:15:30 ; Search time 15 seconds

(without alignments)
365.311 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTRDNLVLVSVSKDGA.....ALMGCKMTATCHCSIHVK 57

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Minimum number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	298	98.0	57	1 NILSA
2	157	51.6	56	1 NIBSA
3	94.5	31.1	52	1 EPSCD
4	91.5	30.1	52	1 EPSCD
5	63	20.7	995	2 S50358
6	59.5	19.6	78	2 C82577
7	59.5	19.6	330	2 T25169
8	59.5	19.6	1574	2 T13954
9	59	19.4	456	2 S20597
10	58.5	19.2	166	2 T33970
11	58.5	19.2	772	2 D56695
12	58	19.1	565	2 A56879
13	57.5	18.9	727	2 A56879
14	57.5	18.9	771	2 S35681
15	56.5	18.6	232	2 S70355
16	56.5	18.6	456	2 A31857
17	56.5	18.6	741	2 B49555
18	56.5	18.6	1700	2 T49527
19	56	18.4	107	2 AF3300
20	56	18.4	506	2 AF3300
21	55.5	18.4	1016	2 T29208
22	55.5	18.3	333	2 T29208
23	55.5	18.3	724	2 T47149
24	55.5	18.3	1620	2 T27283
25	55.5	18.3	1891	2 T13594
26	55.5	18.3	1920	2 T13893
27	55	18.1	267	2 C85075
28	55	18.1	267	2 C85075
29	55	18.1	369	2 S13721

30	55	18.1	374	1 A56436	alcohol dehydrogen
31	55	18.1	384	2 S25771	gsel protein - mou
32	55	18.1	551	2 S23400	bud emergence medl
33	55	18.1	782	2 A10062	conserved hypotet
34	55	18.1	808	1 S62594	replication licens
35	54	17.9	134	2 T22275	hypothetical prote
36	54	17.8	172	2 T50694	transcription fact
37	54	17.8	199	2 T47716	transcription fact
38	54	17.8	343	2 T30233	ornithine cyclodex
39	54	17.8	355	2 C96651	protein T3p18.9 [1
40	54	17.8	511	2 T07787	pyruvate kinase (E
41	54	17.8	518	2 S42091	tyld(56) protein -
42	54	17.8	876	2 B82163	DNA topoisomerase
43	54	17.8	1030	2 B98324	unknown protein, 2
44	54	17.8	1046	2 B98324	probable rnd efflu
45	54	17.8	1046	2 AD2959	AcRb/AcrD/AcrF fam

ALIGNMENTS

RESULT 1

NILSA

nisin precursor - Lactococcus lactis

N:Alternate names: nisin A; nisin Z

C:Species: Lactococcus lactis

C>Date: 21-May-1990 #sequence_revision 12-May-1994 #text_change 18-Jun-1999

C:Accession: A31915; A32809; B45821; A43743; S17858; B48951; S36734; S70485; S16779;

R:Buchanan, G.W.; Banerjee, S.; Hansen, J.N.

J. Biol. Chem. 263, 16260-16266, 1988

A:Title: Structure, expression, and evolution of a gene encoding the precursor of nls

A:Reference number: A92679; MUID:89034093; PMID:3141403

A:Accession: A31915

A:Molecule type: DNA

A:Residues: 1-57 <BUC>

A:Cross-references: GB:J04057; NID:9153816; PIDN:AAA8606.1; PID:9153817

A:Experimental source: ATCC 11454

A>Note: The authors identified the species as Streptococcus lactis

R:Kalella, C.; Ertlan, K.D.

J. Bacteriol. 171, 1597-1601, 1989

A:Title: Nisin, a peptide antibiotic: cloning and sequencing of the nlsA gene and pos

A:Reference number: A32809; MUID:89155467; PMID:2493449

A:Accession: A32809

A:Molecule type: DNA

A:Residues: 1-57 <KAL>

A:Cross-references: GB:M24527; NID:9341189; PIDN:AAA26948.1; PID:9530218

J. Gen. Microbiol. 136, 535-566, 1990

A:Title: Analysis of the genetic determinant for production of the peptide antibiotic

A:Reference number: A45821; MUID:90362041; PMID:2118169

A:Accession: B45821

A:Molecule type: DNA

A:Residues: 1-57 <DOD>

A:Cross-references: GB:M2727; NID:9149440; PIDN:AAA25188.1; PID:9149441

R:Steen, M.T.; Chung, Y.J.; Hansen, J.N.

Appl. Environ. Microbiol. 57, 1181-1188, 1991

A:Title: Characterization of the nlsin gene as part of a polyclonistic operon in the

A:Reference number: A43743; MUID:91282469; PMID:1905517

A:Accession: A43743

A:Molecule type: DNA

A:Residues: 1-57 <STE>

A:Cross-references: GB:M65089; NID:9149447; PIDN:AAA73038.1; PID:9149448

A:Experimental source: ATCC 11454

R:Mulders, J.W.M.; Boerigter, I.J.; Rollem, H.S.; Slezzen, R.J.; de Vos, W.M.

Eur. J. Biochem. 201, 581-584, 1991

A:Title: Identification and characterization of the lactibiotic nlsin Z, a natural nl

A:Reference number: S17858; MUID:92037612; PMID:1935953

A:Accession: S17858

A:Molecule type: DNA

A:Residues: 1-49, N', 51-57 <MUL>

A:Cross-references: EMBL:X61144; NID:944046; PIDN:CAA3440.1; PID:944047

A:Experimental source: strain NIZO 22186

A>Note: nlsin Z allelic variant; amino acid composition and structure determination b

R:Engelke, G.; Gutowski-Eckel, Z.; Hammelmann, M.; Entian, K.D.
 Appl. Environ. Microbiol. 58, 3730-3743, 1992
 A:Title: Biosynthesis of the lantibiotic nisin: genomic organization and membrane localization
 A:Reference number: A48951; MUID:93128945; PMID:1482192
 A:Accession: B48951
 A:Molecule type: DNA
 A:Residues: 1-57 <END>
 A:Cross-references: GB:J68307; NID:944040; PIDN:CAA48380.1; PID:944042
 A:Experimental source: strain 6F3
 A:Note: sequence extracted from NCBI backbone (NCBI:122292, NCBI:122295)
 R:Kulpers, O.P.; Beertshuyzen, M.M.; Slezzen, R.J.; de Vos, W.M.
 Eur. J. Biochem. 216, 281-291, 1993
 A:Title: Characterization of the nisin gene cluster nisABPCIR of *Lactococcus lactis*.
 A:Reference number: S36734; MUID:9373937; PMID:7689965
 A:Accession: S36734
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-57 <KUN>
 A:Cross-references: GB:J16226; NID:9400365; PIDN:AAA25189.1; PID:9400366
 R:Gross, E.; Morell, J.L.
 Am. Chem. Soc. 93, 4634-4635, 1971
 A:Title: The structure of nisin.
 A:Reference number: A54460; MUID:72072901; PMID:5131162
 A:Contents: annotation
 R:Kulpers, O.P.; Rollema, H.S.; de Vos, W.M.; Slezzen, R.J.
 FEBS Lett. 330, 23-27, 1993
 A:Title: Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*, dltA
 A:Reference number: S36142; MUID:93380562; PMID:8370453
 A:Contents: annotation
 R:van der Meer, J.R.; Polman, J.; Beertshuyzen, M.M.; Slezzen, R.J.; Kulpers, O.P.; De Vos
 J. Bacteriol. 175, 2578-2588, 1993
 A:Title: Characterization of the *Lactococcus lactis* nisin A operon genes nisp, encoding
 involved in nisin biosynthesis.
 A:Reference number: A40621; MUID:93239683; PMID:8478324
 A:Contents: annotation
 R:Chan, W.C.; Leyland, M.; Clark, J.; Dodd, H.M.; Lian, L.Y.; Gasson, M.J.; Bycroft, B.W.
 FEBS Lett. 390, 129-132, 1996
 A:Title: Structure-activity relationships in the peptide antibiotic nisin: antibacterial
 A:Reference number: S70485; MUID:96305786; PMID:8706842
 A:Accession: S70485
 A:Molecule type: protein
 A:Residues: 24-57 <CHN>
 C:Comment: Nisin is secreted as an inactive precursor and then activated by cleavage with
 C:Genetics:
 A:Gene: span; nisa
 C:Superfamily: subtilin precursor
 C:Keywords: antibiotic; lantibiotic
 F:2-23/Domain: propeptide #status experimental <PRO>
 C:Keywords: antibiotic; lantibiotic
 F:2-24/Domain: propeptide #status experimental <MAT>
 F:3-24/Cleavage site: Arg-11e (protease nisp) #status experimental
 F:3-25/Modified site: dehydrobutyryne (Thr) #status experimental
 F:26-30/Cross-link: sn-(25,6r)-lanthionine (Ser-Cys) #status experimental
 F:26-30/Cross-link: sn-(25,6r)-lanthionine (Ser-Cys) #status experimental
 F:31-34/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:36-42/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:46-49/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:48-51/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:56/Modified site: dehydroalanine (Ser) #status experimental

Query Match 98.0%; Score 298; DB 1; Length 57;
 Best local similarity 98.2%; Pred. No. 6; 3e-27;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MSTRDENLDVSVKSGSPRITSICCTPGCKTGALMGCMKATATGCHGSIHVS 57
 DB 1 MSTRDENLDVSVKSGSPRITSICCTPGCKTGALMGCMKATATGCHGSIHVS 57

RESULT 2
 subtilin precursor - *Bacillus subtilis* (strain ATCC 6633)
 N:Alternate names: Spas; subtilin A
 N:Contains: subtilin B

C:Species: *Bacillus subtilis*
 C:Date: 30-Sep-1989 #sequence-revision 12-May-1994 #text-change 21-Jul-2000
 C:Accession: A28112; D42655; D43935; I40514; I39980; S36142
 R:Banerjee, S.; Hansen, J.N.
 J. Biol. Chem. 263, 9508-9514, 1988
 A:Title: Structure and expression of a gene encoding the precursor of subtilin, a sma
 A:Reference number: A28112; MUID:88243844; PMID:2837490
 A:Accession: A28112
 A:Molecule type: DNA
 A:Residues: 1-56 <BN>
 A:Cross-references: GB:J03767; NID:9143718; PIDN:AAA22841.1; PID:9143719
 R:Chung, Y.J.; Steen, M.T.; Hansen, J.N.
 J. Bacteriol. 174, 1417-1422, 1992
 A:Title: The subtilin gene of *Bacillus subtilis* ATCC 6633 is encoded in an operon the
 A:Reference number: A42655; MUID:92138640; PMID:1735728
 A:Accession: D42655
 A:Molecule type: DNA
 A:Residues: 1-56 <CHD>
 A:Cross-references: GB:M83944; NID:9143557; PIDN:AAA22772.1; PID:9143561
 A:Experimental source: ATCC 6633
 A:Note: sequence extracted from NCBI backbone (NCBI:79670)
 R:Klein, C.; Kaletta, C.; Schnell, N.; Entian, K.D.
 Appl. Environ. Microbiol. 58, 132-142, 1992
 A:Title: Analysis of genes involved in biosynthesis of the lantibiotic subtilin.
 A:Reference number: A43935; MUID:92171481; PMID:1539969
 A:Accession: D43935
 A:Molecule type: DNA
 A:Residues: 1-56 <KLE1>
 A:Cross-references: GB:M86869; NID:9143713; PIDN:AAA22840.1; PID:9143717
 A:Experimental source: ATCC 6633
 A:Note: sequence extracted from NCBI backbone (NCBI:84011, NCBI:84020)
 R:Klein, C.; Entian, K.D.
 Appl. Environ. Microbiol. 60, 2793-2801, 1994
 A:Title: Genes involved in self-protection against the lantibiotic subtilin produced
 A:Reference number: I40511; MUID:94368094; PMID:8085823
 A:Accession: I40511
 A:Status: preliminary
 A:Molecule type: translated from GB/EMBL/DBJ
 A:Residues: 1-56 <KLE2>
 A:Cross-references: EMBL:U09819; NID:92702240; PIDN:AB91589.1; PID:9595319
 A:Experimental source: ATCC 6633
 R:Chan, W.C.; Bycroft, B.W.; Leyland, M.L.; Lian, L.Y.; Yang, J.C.; Roberts, G.C.K.
 FEBS Lett. 300, 56-62, 1992
 A:Title: Sequence-specific resonance assignment and conformational analysis of subtil
 A:Reference number: A4571; MUID:92192284; PMID:1547888
 A:Contents: annotation
 R:Chan, W.C.; Bycroft, B.W.; Leyland, M.L.; Lian, L.Y.; Roberts, G.C.K.
 Biochem. J. 291, 23-27, 1993
 A:Title: A novel post-translational modification of the peptide antibiotic subtilin:
 A:Reference number: A53265; MUID:93228611; PMID:8471040
 R:Kulpers, O.P.; Rollema, H.S.; de Vos, W.M.; Slezzen, R.J.
 FEBS Lett. 330, 23-27, 1993
 A:Title: Biosynthesis and secretion of a precursor of nisin Z by *Lactococcus lactis*,
 A:Reference number: S36142; MUID:93380562; PMID:8370453
 A:Contents: annotation
 C:Comment: *Bacillus subtilis* strain ATCC 6633 carries this gene that is not found in
 C:Genetics:
 A:Gene: spas
 C:Superfamily: subtilin precursor
 C:Keywords: antibiotic; blocked amino end; lantibiotic
 F:2-24/Domain: propeptide #status predicted <SIG>
 F:25-56/Product: subtilin A #status experimental <MATB>
 F:25-56/Product: subtilin B #status experimental <MATB>
 F:25-56/Product: subtilin C #status experimental <MATB>
 F:27-31/Cross-link: sn-(25,6r)-lanthionine (Ser-Cys) #status experimental
 F:27-31/Cross-link: sn-(25,6r)-lanthionine (Ser-Cys) #status experimental
 F:32-35/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:37-43/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:42/Modified site: (2)-dehydrobutyryne (Thr) #status experimental
 F:47-50/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:49-52/Cross-link: (25,35,6r)-3-methyl-lanthionine (Thr-Cys) #status experimental
 F:55/Modified site: dehydroalanine (Ser) #status experimental

C:species: *Xylella fastidiosa*
C:date: 18-Aug-2000 #sequence_revision 20-Aug-2000
C:accession: C82577
R:anonymous, The Xylella fastidiosa Consortium of the Organoleide Sequ
ature 406, 151-157, 2000

Db 19 DTASVSGESTVAPRRKQIYLYLGNFNPQYLSLTPCNSGCSST-----CNCNTRATCSTS 73

RESULT 11

D56695
transducin-like enhancer of split homolog TLE-3 - human
C/Species: Homo sapiens (man)
C/Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 26-May-2000
C/Accession: D56695
R/Stifani, S.; Blauweller, C.M.; Redhead, N.J.; Hill, R.E.; Artavanis-Tsakonas, S.
Nucleic Acids Res. 23, 119-127, 1995
A/Title: Human homologs of a Drosophila enhancer of split gene product define a novel family of transcription factors
A/Reference number: A56695; MUID:93265135; PMID:1303260
A/Accession: D56695
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-772 <STI>
A/Cross-references: GB:M99438; NID:g307513; PIDN:AAA61194.1; PID:g307514
A/Genetics:
Gene: GDB:TLE3; ESG: ESG3
Cross-references: GDB:228049; OMIM:600190
A/Map position: 15pter-15qter
C/Superfamily: unassigned WD repeat proteins; WD repeat homology
C/Keywords: nucleus
F:482-514/Domain: WD repeat homology <WD1>
F:528-561/Domain: WD repeat homology <WD2>
F:614-647/Domain: WD repeat homology <WD3>
F:696-729/Domain: WD repeat homology <WD4>
F:734-770/Domain: WD repeat homology <WD5>

Query Match

Best Local Similarity 19.2%; Score 58.5; DB 2; Length 772;
Best Local Similarity 38.3%; Pred. No. 70;
Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

QY 2 STKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTAT 48
Db 267 SPENGIDKARKSLKADAPSPASVASSSTSPSSKTKDL-GHNDKSSST 312

RESULT 12

F84721
probable RING zinc finger protein (Imported) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: F84721
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
Wang, K.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Muller, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nucleic Acids Res. 29, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: F84721
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-565 <STO>
A/Cross-references: GB:AEO02093; NID:g4582446; PIDN:AAD24830.1; GSPDB:GND00139
A/Genetics:
A:Gene: Atg31510
A/Map position: 2

Query Match 19.1%; Score 58; DB 2; Length 565;
Best Local Similarity 28.8%; Pred. No. 61;
Matches 23; Conservative 9; Mismatches 20; Indels 28; Gaps 4;

QY 3 TNDPNDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTAT 38
Db 426 SKDFNDFRTKAGLTSVTKNYPENLVKALENGLADVDHAAKSSKSTSSKSTGSCSKTPE 485

QY 39 LM-----GCNMKTATCHCSIH 54
Db 486 LVDPDGGTKPELIVCNCTRH 505

RESULT 13

A56879
diacylglycerol kinase (EC 2.7.1.107) alpha - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 18-Jun-1999
C/Accession: A56879
R/Goto, K.; Matsubara, M.; Kondo, H.; Yusa, H.; Sakane, F.; Kanoh, H.
Brain Res. Mol. Brain Res. 16, 75-87, 1992
A/Title: Gene cloning, sequence, expression and in situ localization of 80 kDa diacyl
A/Reference number: A56879; MUID:93095720; PMID:1339302
A/Contents: brain
A/Accession: A56879
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-727 <GOT>
A/Cross-references: GB:S49760; NID:g261423; PIDN:AAB24434.1; PID:g261424
A/Note: sequence extracted from NCBI backbone (NCBI:120083, NCBI:120084)
C/Superfamily: human diacylglycerol kinase; calmodulin repeat homology; protein kinas
C/Keywords: ATP; calcium binding; duplication; EF hand; phosphotransferase; zinc
F:108-140/Domain: calmodulin repeat homology <EF1>
F:153-185/Domain: calmodulin repeat homology <EF2>
F:204-251/Domain: protein kinase C zinc-binding repeat homology <K21>
F:268-317/Domain: protein kinase C zinc-binding repeat homology <K22>

Query Match

Best Local Similarity 18.9%; Score 57.5; DB 2; Length 727;
Best Local Similarity 24.6%; Pred. No. 86;
Matches 16; Conservative 6; Mismatches 28; Indels 15; Gaps 2;

QY 1 MSTRKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTAT-----C 49
Db 246 MKAPCEVSTYAKSRDVGQPHWVRG-----GCHSGRDRDROCKKIRITHSLGLHCWVC 301

QY 50 HCSIH 54
Db 302 HLEIH 306

RESULT 14

S35681
ESG protein - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-May-2000
C/Accession: S35681; S34162
R/Miyasaka, H.; Choudhury, B.K.; Hou, E.W.; Li, S.S.L.
Eur. J. Biochem. 216, 343-352, 1993
A/Title: Molecular cloning and expression of mouse and human cDNA encoding AES and ES
A/Reference number: S35678; MUID:93373944; PMID:8365415
A/Accession: S35681
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-771 <MTY>
A/Cross-references: EMBL:X73360; NID:g313235; PIDN:CAA51770.1; PID:g313236
A/Superfamily: unassigned WD repeat proteins; WD repeat homology
C/Keywords: phosphoprotein
F:527-560/Domain: WD repeat homology <WD1>
F:527-560/Domain: WD repeat homology <WD2>
F:613-646/Domain: WD repeat homology <WD3>
F:695-728/Domain: WD repeat homology <WD4>
F:736-769/Domain: WD repeat homology <WD5>

Query Match 18.9%; Score 57.5; DB 2; Length 771;
Best Local Similarity 38.3%; Pred. No. 91;
Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

QY 2 STKDFNLDLVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTAT 48
Db 266 SPENGIDKARKSLKADAPSPASVASSSTSPSSKTKDL-GHNDKSSST 311

RESULT 15

S70355
phosphoribosylanthranilate isomerase (EC 5.3.1.24) - Lipomyces starkeyi
C/Species: Lipomyces starkeyi
C/Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S70355
R:Biggell, G.R.; Bruce, I.J.; Evans, I.H.
Curr. Genet. 30, 83-88, 1996
A:Title: Electrophoretic karyotype of the amyloidic Yeast *Lipomyces starkeyi* and clonin
A:Reference number: S70355; MUID:96269934; PMID:8662214
A:Accession: S70355
A:Molecule type: DNA
A:Residues: 1-232 <BIG>
A:Cross-references: EMBL:Z68292; NID:q1134847; PIDN:CAA92584.1; PID:q1161576
A:Note: the authors translated the codon GCC for residue 211 as Glu
C:Genetics:
A:Gene: TRP1
C:Superfamily: phosphoribosylanthranilate isomerase; trpF homology
C:Keywords: Intramolecular oxidoreductase; isomerase; tryptophan biosynthesis
F:14-228/Domain: trpF homology <TRF>

Query Match 18.6%; Score 56.5; DB 2; Length 232;
Best Local Similarity 36.7%; Pred. No. 41;
Matches 11; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

24 ITSTSLCTPGCKTGALMGCMKMTATCHCSI 53
:::||||| | : | : |
3 VSTSLCTPIVKI-----CGLTVEAHACAI 27

Search completed: June 7, 2003, 15:19:26
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:12:29 ; Search time 11 Seconds
(without alignments)
214.923 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304
Sequence: 1 MSTKDFNLDLVSVSKDSCA.....ALMGCMKATCCHSIHVK 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

1 number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwisProt_40.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298	98.0	57	1	PI1068 lactococcus
2	291	95.7	57	1	LANN_LACLA
3	157	51.6	56	1	LANZ_LACLA
4	94.5	31.1	52	1	LANE_BACSU
5	91.5	30.1	52	1	LANE_STARP
6	72	23.7	46	1	LANG-STAGA
7	63	20.7	46	1	SRVA-STREX
8	61	20.1	45	1	YF48_HUMAN
9	59	19.4	45	1	YF48_HUMAN
10	58.5	19.2	44	1	YF48_HUMAN
11	58.5	19.2	44	1	YF48_HUMAN
12	57.5	18.9	44	1	YF48_HUMAN
13	57.5	18.9	44	1	YF48_HUMAN
14	57.5	18.9	44	1	YF48_HUMAN
15	56.5	18.6	43	1	YF48_HUMAN
16	56.5	18.6	43	1	YF48_HUMAN
17	56.5	18.6	43	1	YF48_HUMAN
18	56.5	18.6	43	1	YF48_HUMAN
19	56.5	18.6	43	1	YF48_HUMAN
20	56.5	18.6	43	1	YF48_HUMAN
21	56.5	18.6	43	1	YF48_HUMAN
22	56.5	18.6	43	1	YF48_HUMAN
23	56.5	18.6	43	1	YF48_HUMAN
24	56.5	18.6	43	1	YF48_HUMAN
25	56.5	18.6	43	1	YF48_HUMAN
26	56.5	18.6	43	1	YF48_HUMAN
27	56.5	18.6	43	1	YF48_HUMAN
28	56.5	18.6	43	1	YF48_HUMAN
29	56.5	18.6	43	1	YF48_HUMAN
30	56.5	18.6	43	1	YF48_HUMAN
31	56.5	18.6	43	1	YF48_HUMAN
32	56.5	18.6	43	1	YF48_HUMAN
33	56.5	18.6	43	1	YF48_HUMAN

34	54	17.8	511	1	KPYC-SOYB	042806 glycine max
35	54	17.8	808	1	MCM3_HUMAN	P25205 homo sapien
36	54	17.8	876	1	TOP1_VIBCH	O9KRB2 vibrio chol
37	54	17.8	1257	1	RBB1_HUMAN	P29374 homo sapien
38	54	17.8	2193	1	POLG_CXAL6	O65900 c genome po
39	53.5	17.6	166	1	THI2_HUMAN	O99757 homo sapien
40	53.5	17.6	442	1	TIC-CHLTR	O84713 chlamydia t
41	53.5	17.6	3672	1	LM12_CAEEL	Q21313 caenorhabd1
42	53	17.4	298	1	YIE4_YEAST	P40529 saccharomyc
43	53	17.4	332	1	MYOD_DROME	P22816 drosophila
44	53	17.4	374	1	ADH7_RAT	P41682 rattus norv
45	53	17.4	463	1	GLP1_HUMAN	P43220 homo sapien

ALIGNMENTS

RESULT 1	ID	STANDARD	PRT	57 AA
AC	PI3068;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Lactobacillus nisin A precursor.			
GN	SPAN OR NISA.			
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.			
OX	NCBI_TaxID=1360;			
RM	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 11454 / DSM 20729 / NCTC 496;			
RA	MEDLINE-89034093; PubMed-3141403;			
RT	Buchanan G.W., Banerjee S., Hansen J.N.,			
RT	"Structure, expression, and evolution of a gene encoding the			
RT	precursor of nisin, a small protein antibiotic."			
RT	J. Biol. Chem. 263:16260-16266(1988).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 11454 / DSM 20729 / NCTC 496;			
RA	MEDLINE-91282469; PubMed-1905517;			
RT	Steen M.T., Chung Y.J., Hansen J.N.,			
RT	"Characterization of the nisin gene as part of a polycistronic operon			
RT	in the chromosome of Lactococcus lactis ATCC 11454."			
RT	Appl. Environ. Microbiol. 57:1181-1188(1991).			
RM	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-6F3;			
RA	MEDLINE-89155467; PubMed-2493449;			
RT	Kaletta C., Entian K.-D.,			
RT	"Nisin, a peptide antibiotic: cloning and sequencing of the nisa gene			
RT	and posttranslational processing of its peptide product."			
RT	J. Bacteriol. 171:1597-1601(1989).			
RM	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-6F3;			
RA	MEDLINE-91282469; PubMed-1482192;			
RT	Engelke G., Gutowski-Eckel Z., Hammelmann M., Entian K.-D.,			
RT	"Biosynthesis of the lantibiotic nisin: genomic organization and			
RT	membrane localization of the nisa protein."			
RT	Appl. Environ. Microbiol. 58:3730-3743(1992).			
RM	[5]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NIZO R5;			
RA	MEDLINE-93373937; PubMed-7689965;			
RT	Kuipers O.P., Beerhuizen M.M., Steen R.J., de Vos W.M.,			
RT	"Characterization of the nisin gene cluster nisaRTIPR of Lactococcus			
RT	lactis. Requirement of expression of the nisa and nisa genes for			
RT	development of immunity."			
RT	Eur. J. Biochem. 216:281-291(1993).			
RM	[6]			
RP	SEQUENCE OF 24-57.			
RA	Gross E.,			

RL (in) Friedman M. (eds.);
 RL Protein cross-linking, pp.131-153, Plenum Press, New York (1977).
 RN [7]
 RP SEQUENCE OF 24-57.
 RX MEDLINE=72072901; PubMed=5111162;
 RA Gross E., Morell J.L.;
 RT "The structure of nisin";
 RL J. Am. Chem. Soc. 93:4634-4635(1971).
 RN [8]
 RP STRUCTURE BY NMR.
 RX MEDLINE=92111494; PubMed=1765078;
 RA van de Ven F.J., van den Hooven H.W., Konings R.N.H., Hilbers C.W.;
 RT "NMR studies of lantibiotics. The structure of nisin in aqueous solution";
 RL Eur. J. Biochem. 202:1181-1188(1991).
 RN [9]
 RP STRUCTURE BY NMR.
 RX MEDLINE=92246867; PubMed=1575686;
 RA Llan L.-Y., Chan W.C., Morley S.D., Roberts G.C.K., Bycroft B.W., Jackson D.;
 RT "Solution structures of nisin A and its two major degradation products determined by NMR";
 RL Biochem. J. 283:413-420(1992).
 RN [10]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93202265; PubMed=8454055;
 RA van den Hooven H.W., Fogolari F., Rolletta H.S., Konings R.N.H., Hilbers C.W., van de Ven F.J.;
 RT "NMR and circular dichroism studies of the lantibiotic nisin in non-aqueous environments";
 RL FEBS Lett. 319:189-194(1993).
 RN [11]
 RP STRUCTURE BY NMR.
 RX MEDLINE=93120109; PubMed=8418850;
 RA Saller M., Helms G.L., Henkel T., Niemczura W.P., Stiles M.E., Vederas J.C.;
 RT "15N- and 13C-labeled media from Anaerobaculum sp. for universal isotopic labeling of bacteriophages: NMR resonance assignments of leucocin A from Leuconostoc gelidium and nisin A from Lactococcus lactis";
 RL Biochemistry 32:310-318(1993).
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC) ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF LANTIBIOTICS IS BASED ON DEPOLARIZATION OF EMERGIZED BACTERIAL CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS TRANSMEMBRANE PORES.
 CC -1- PROM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND CLEAVAGE OF THE MODIFIED PRECURSOR.
 CC -1- MISCELLANEOUS: USED AS A FOOD PRESERVATIVE.
 CC -1- MISCELLANEOUS: THE NISA GENE IS FOUND BOTH ON CHROMOSOMAL AND PLASMID DNA. THE SEQUENCES REPORTED ARE ABSOLUTELY IDENTICAL.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J04057; AAA86606.1;
 DR EMBL; M65089; AAA73038.1;
 DR EMBL; M24527; AAA26948.1;
 DR EMBL; X68307; CAA48380.1;
 DR EMBL; M27277; AAA25188.1;
 DR EMBL; D00696; BAA00602.1;
 DR EMBL; L16226; AAA25189.1;
 DR EMBL; M79445; AAA25198.1;
 DR PIR; A31915; NILLSA.
 DR PIR; B48951; B48951.
 DR PIR; S36734; S36734.

DR InterPro: IPR001049; Gallidermin.
 DR InterPro: IPR000446; Nisin.
 DR Pfam: PF02052; Gallidermin; 1.
 DR PRINTS: PR00324; NISIN.
 KW Antibiotic; Bacteriocin; Lantibiotic; Plasmid.
 FT PROPEP 1 23
 FT CHAIN 24 57
 FT MOD_RES 25 25
 FT MOD_RES 26 26
 FT MOD_RES 28 28
 FT MOD_RES 31 31
 FT MOD_RES 36 36
 FT MOD_RES 46 46
 FT MOD_RES 48 48
 FT MOD_RES 56 56
 FT MOD_RES 56 56
 FT THIOETH 31 34
 FT THIOETH 31 34
 FT THIOETH 36 42
 FT THIOETH 46 49
 FT THIOETH 48 51
 SO SEQUENCE 57 AA; 5963 MW; 315E4428AC70BFBA CRC64;
 Query Match 98.0%; Score 298; DB 1; Length 57;
 Best Local Similarity 98.2%; Pred No. 2e-30;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 MSTRKDFNLVSVSKKSGASPRITSTSLCTPGCKTGALMGCMKTAHCISHWK 57
 Db 1 MSTRKDFNLVSVSKKSGASPRITSTSLCTPGCKTGALMGCMKTAHCISHWK 57
 RESULT 2
 ID LANTZ_LACIA STANDARD; PRT; 57 AA.
 AC P29559;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lantibiotic nisin Z precursor.
 GN Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OC NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIZO 22186;
 RX MEDLINE=92037612; PubMed=1935953;
 RA Mulders J.W.M., Boerrigter I.J., Rolletta H.S., Siezen R.J., de Vos W.M.;
 RT "Identification and characterization of the lantibiotic nisin Z, a natural nisin variant";
 RL Eur. J. Biochem. 201:581-584(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 7638;
 RA Araya T., Ishibashi N., Shimamura S.;
 RT "Genetic evidence that Lactococcus lactis JCM7638 produces a mutated form of nisin";
 RL J. Gen. Appl. Microbiol. 38:271-278(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N8;
 RX MEDLINE=95352820; PubMed=7626780;
 RA Immonen T., Ye S., Ra R., Qiao M., Paulin L., Sarris P.E.J.;
 RT "The codon usage of the nisz operon in Lactococcus lactis N8 suggests a non-lactococcal origin of the conjugative nisin-sucrose transposon";
 RL DNA Seq. 5:203-218(1995).
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC) ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF LANTIBIOTICS IS BASED ON DEPOLARIZATION OF EMERGIZED BACTERIAL CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS TRANSMEMBRANE PORES.

CC -1- P.T.M: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND
 CC CLEAVAGE OF THE MODIFIED PRECURSOR.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X61144; CAA43440.1; -
 CC EMBL: D10768; BAA01598.1; -
 CC EMBL: 218947; CAA79467.1; -
 CC PIR: A31915; NILSA.
 CC InterPro: IPR001049; Galildermln.
 CC InterPro: IPR000446; Nisin.
 CC Pfam: PF02052; Galildermln; 1.
 CC PRINTS: PR00324; Nisin.
 CC Antibiotic; Bacteriocin; Lantibiotic.
 CC PROPEP 1 23
 CC CHAIN 24 57
 CC MOD_RES 25 25 LANTIBIOTIC NISIN Z.
 CC MOD_RES 26 26 DHB (2,3-DIDEHYDROBUTYRINE).
 CC MOD_RES 28 28 D-ALANINE.
 CC MOD_RES 31 31 DHA (2,3-DIDEHYDROALANINE).
 CC MOD_RES 36 36 D-ABU (AMINOBOUYRIC ACID).
 CC MOD_RES 46 36 D-ABU (AMINOBOUYRIC ACID).
 CC MOD_RES 48 48 D-ABU (AMINOBOUYRIC ACID).
 CC MOD_RES 56 56 DHA (2,3-DIDEHYDROALANINE).
 CC THIOETH 31 34 ALA-S-CYS (LANTHIONINE).
 CC THIOETH 36 42 ABU-S-CYS (BETA-METHYLLANTHIONINE).
 CC THIOETH 46 49 ABU-S-CYS (BETA-METHYLLANTHIONINE).
 CC THIOETH 48 51 ABU-S-CYS (BETA-METHYLLANTHIONINE).
 CC THIOETH 50 50 N -> H (IN STRAIN JCM7638).
 CC VARIANT 50 50 DF5E4428AC70BEEF CRC64;
 CC SEQUENCE 57 AA: 5940 MW: 6554428AC70BEEF CRC64;
 CC
 CC Query Match Best Local Similarity 95.7% Score 291; DB 1; Length 57;
 CC Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 1 MSTRDPLNDLVSVKSGASPRITSTSLCTPGCKTGALMGCKNTKATCHCSHVSK 57
 CC 1 MSTRDPLNDLVSVKSGASPRITSTSLCTPGCKTGALMGCKNTKATCHCSHVSK 57
 CC
 CC RESULT 3
 CC LANS_BACSU STANDARD; PRT; 56 AA.
 CC AC P10946;
 CC DT 01-JUL-1989 (Rel. 11, Created)
 CC DT 01-JUL-1989 (Rel. 11, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Lantibiotic subtilin precursor.
 CC GN SPAS OR SUB.
 CC OS Bacillus subtilis.
 CC OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC OX NCBI_TaxID=1423;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=88243844; PubMed=2837490;
 CC RA Banerjee S., Hansen J.N.;
 CC RT "Structure and expression of a gene encoding the precursor of
 CC subtilin, a small protein antilabiotic";
 CC RL J Biol. Chem. 263:9508-9514(1988).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN-ATCC 6633 / LH45;
 CC MEDLINE=92138640; PubMed=1735728;

RA Chung Y.J., Steen M.T., Hansen J.N.;
 RT "The subtilin gene of Bacillus subtilis ATCC 6633 is encoded in an
 RT operon that contains a homolog of the hemolysin B transport
 RT protein.";
 RT J. Bacteriol. 174:1417-1422(1992).
 RN [3]
 RN RP SEQUENCE FROM N.A.
 RN RC STRAIN-ATCC 6633 / LH45;
 RN RX MEDLINE=92171481; PubMed=1539969;
 RN RA Klein C., Kaletta C., Schnell N., Entian K.-D.;
 RN RT "Analysis of genes involved in biosynthesis of the lantibiotic
 RN subtilin.";
 RN RL Appl. Environ. Microbiol. 58:132-142(1992).
 RN [4]
 RN RP SEQUENCE OF 25-56.
 RN RX MEDLINE=75040028; PubMed=4154277;
 RN RA Gross E., Kiltz H.H., Nebelin E.;
 RN RT "Subtilin, VI: the structure of subtilin.";
 RN RL Hoppe-Seyler's Z. Physiol. Chem. 354:810-812(1973).
 RN [5]
 RN RP MODE OF ACTION.
 RN RX MEDLINE=89276381; PubMed=2471644;
 RN RA Schueller F., Benz R., Sahl H.-G.;
 RN RT "The peptide antibiotic subtilin acts by formation of
 RN voltage-dependent multi-state pores in bacterial and artificial
 RN membranes.";
 RN RL Eur. J. Biochem. 182:181-186(1989).
 RN [6]
 RN RP STRUCTURE BY NMR.
 RN RC STRAIN-ATCC 6633 / LH45;
 RN RX MEDLINE=92192284; PubMed=1547888;
 RN RA Chan W.C., Bycroft B.W., Leylands M.L., Llan L.-Y., Yang J.C.,
 RN RA Roberts G.C.K.;
 RN RT "Sequence-specific resonance assignment and conformational analysis
 RN of subtilin by 2D NMR.";
 RN FEBS Lett. 300:56-62(1992).
 RN [7]
 RN RP MUTAGENESIS OF SER-29.
 RN RX MEDLINE=93167833; PubMed=8434932;
 RN RA Liu W., Hansen J.N.;
 RN RT "The antimicrobial effect of a structural variant of subtilin against
 RN outgrowing Bacillus cereus T spores and vegetative cells occurs by
 RN different mechanisms.";
 RN RL Appl. Environ. Microbiol. 59:648-651(1993).
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC)
 CC LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL
 CC CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS
 CC TRANSMEMBRANE PORES.
 CC -1- P.T.M: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF
 CC THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE
 CC BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND
 CC CLEAVAGE OF THE MODIFIED PRECURSOR.
 CC -1- MISCELLANEOUS: SUBTILIN ACTIVITY IS OBSERVED DURING STATIONARY
 CC PHASE, BUT NOT DURING EXPONENTIAL GROWTH.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER LANTIBIOTICS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J03767; AAA22841.1; -
 CC EMBL: M86869; AAA22840.1; -
 CC EMBL: M83944; AAA22772.1; -
 CC EMBL: M99263; AAA22778.1; -
 CC EMBL: U09819; AAB91589.1; -
 CC PIR: A28112; NTBSSA.
 CC InterPro: IPR001049; Galildermln.
 CC InterPro: IPR000446; Nisin.

DR Pfam: PF02052; Gallidermin; 1.
 DR PRINTS: PR00324; NISIN.
 KM Antibiotic; Bacteriocin; Lantibiotic.
 FT PROPEP 1 24
 FT CHAIN 25 56
 FT MOD_RES 27 27
 FT MOD_RES 29 29
 FT MOD_RES 32 32
 FT MOD_RES 37 37
 FT MOD_RES 42 42
 FT MOD_RES 47 47
 FT MOD_RES 49 49
 FT MOD_RES 55 55
 FT THIOETH 27 31
 FT THIOETH 32 35
 FT THIOETH 37 43
 FT THIOETH 47 50
 FT THIOETH 49 52
 FT MUTAGEN 29 29
 SEQUENCE 56 AA; 6218 MW; DA9707FBFA1EBBA CRC64;
 Query Match 51.6%; Score 157; DB 1; Length 56;
 Best Local Similarity 59.2%; Pred. No. 5.5e-13;
 Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
 OY 5 FNLDIVSVSKKSGASPRITSTSLCTPGCKTGALMGCKMTATCHRSI 53
 DB 6 DFDDIVVKSVDKSKITPPWKSESLCTPGCVTALQTCFLQTLQCNCKI 54
 RESULT 4
 LANE_STAGP STANDARD; PRT; 52 AA.
 ID LANE_STAGP STANDARD; PRT; 52 AA.
 AC P08136; Q54093;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lantibiotic epidermin precursor.
 GN EPIA.
 OS Staphylococcus epidermidis.
 OG Plasmid ptn 32.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TU 3298 / DSM 3095;
 RC MEDLINE=88216821; PubMed=2835685;
 RX Schnell N., Entian K.-D., Schneider U., Gotz F., Zahner H., Kellner R., Jung G.;
 "Prepeptide sequence of epidermin, a ribosomally synthesized antibiotic with four sulphide-rings.";
 RT Nature 333:276-278(1988).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-TU 3298 / DSM 3095;
 RC MEDLINE=92155237; PubMed=1740156;
 RX Schnell N., Engelke G., Augustin J., Rosenstein R., Ungermann V., Goetz F., Entian K.-D.;
 "Analysis of genes involved in the biosynthesis of lantibiotic epidermin.";
 RT Eur. J. Biochem. 204:57-68(1992).
 CC -1- FUNCTION: LANTHIONINE-CONTAINING PEPTIDE ANTIBIOTIC (LANTIBIOTIC) ACTIVE ON GRAM-POSITIVE BACTERIA. THE BACTERICIDAL ACTIVITY OF LANTIBIOTICS IS BASED ON DEPOLARIZATION OF ENERGIZED BACTERIAL CYTOPLASMIC MEMBRANES, INITIATED BY THE FORMATION OF AQUEOUS TRANSMEMBRANE PORES.
 CC -1- PTM: MATURATION OF LANTIBIOTICS INVOLVES THE ENZYMIC CONVERSION OF THR, SER, AND CYS INTO DEHYDRATED AA AND THE FORMATION OF SULFIDE BRIDGES. THIS IS FOLLOWED BY THE MEMBRANE TRANSLOCATION AND CLEAVAGE OF THE MODIFIED PRECURSOR.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY TO OTHER TYPE A LANTIBIOTICS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC
 CC EMBL: X07840; CA30689.1; -
 CC EMBL: X07840; CA30690.1; -
 CC EMBL: X62386; CA44252.1; -
 CC EMBL: A12927; CA401070.1; -
 CC PIR: S00768; EPSED.
 DR InterPro: IPR001049; Gallidermin.
 DR Pfam: PF02052; Gallidermin; 1.
 DR PRINTS: PR00323; GALLIDERMIN.
 KM Antibiotic; Bacteriocin; Lantibiotic; Plasmid.
 FT PROPEP 1 30
 FT CHAIN 31 52
 FT MOD_RES 33 33
 FT MOD_RES 38 38
 FT MOD_RES 44 44
 FT MOD_RES 46 46
 FT MOD_RES 49 49
 FT MOD_RES 52 52
 FT MOD_RES 52 52
 FT THIOETH 33 37
 FT THIOETH 38 41
 FT THIOETH 46 51
 FT THIOETH 49 52
 SEQUENCE 52 AA; 5632 MW; 8B1AD2875BF10D6D CRC64;
 Query Match 31.1%; Score 94.5; DB 1; Length 52;
 Best Local Similarity 62.2%; Pred. No. 2.7e-05;
 Matches 23; Conservative 3; Mismatches 6; Indels 5; Gaps 3;
 OY 6 FNLDIVSVSKK---DSGASPRITSTSLCTPGC-KTGA 38
 DB 11 FNLDVVKNSDSDGAPRIASKFLCTPGCAKTGS 46
 RESULT 5
 LANG_STAGA STANDARD; PRT; 52 AA.
 ID LANG_STAGA STANDARD; PRT; 52 AA.
 AC P21838;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lantibiotic gallidermin precursor.
 GN GDMA.
 OS Staphylococcus gallinarum.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=1293;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=89305540; PubMed=2765032;
 RX Schnell N., Entian K.-D., Goetz F., Hoerner T., Kellner R., Jung G.;
 "Structural gene isolation and prepeptide sequence of gallidermin, a new lanthionine containing antibiotic.";
 RT FEWS Microbiol. Lett. 49:263-267(1989).
 RL [2]
 RN SEQUENCE OF 31-52.
 RP STRAIN-TU 3928;
 RC MEDLINE=89030695; PubMed=3181159;
 RX Kellner R., Jung G., Hoerner T., Zaehner H., Schnell N., Entian K.-D., Goetz F.;
 "Gallidermin: a new lanthionine-containing polypeptide antibiotic.";
 RT Eur. J. Biochem. 177:53-59(1988).
 CC [3]
 CC STRUCTURE BY NMR.
 CC MEDLINE=92032577; PubMed=1932575;
 CC Freund S., Jung G., Gutbrod O., Folkers G., Gibbons W.A., Allgaier H., Werner R.;
 "The solution structure of the lantibiotic gallidermin.";

Db 676 TTDSNGNVYITTTVPCSSSTATITSCDETGTCHVSTGAVTVETVSKSYTAIVTHCD 735
 QY 41 --GCNMKTATCCHCSIHVS 56
 Db 736 DNGCNKTKVTYSECKETS 753

RESULT 8

YF48_HUMAN STANDARD; PRT; 732 AA.
 ID YF48_HUMAN
 AC Q9HGM4; Q9H975; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein KIAA1548.
 GN KIAA1548.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CC NCBI_TaxID=9606;

(1)

SEQUENCE FROM N.A.
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Makamatsu A.,
 RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.,
 RA "NDO human cDNA sequencing project";
 RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RM SEQUENCE OF 188-732 FROM N.A.
 RP TISSUE=Brain;
 RC MEDLINE=20450683; PubMed=10997877;
 RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes.
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which
 RT code for large proteins in vitro.";
 RL DNA Res. 7:273-281(2000).

-1- SIMILARITY: CONTAINS 1 BAND 4.1-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 CC DR EMBL; AK023019; BAB14360.1; -
 CC EMBL; AB046768; BAB13374.1; -
 CC InterPro; IPR000299; Band_4.1.
 CC Pfam; PF00373; Band_41; 1.
 CC PRINTS; PR000935; BAND41.
 CC DR SMART; SM00295; B41; 1.
 CC DR PROSITE; PS00660; BAND_41_1; 1.
 CC DR PROSITE; PS00661; BAND_41_2; 1.
 CC DR PROSITE; PS00662; BAND_41_3; 1.
 CC DR PROSITE; PS00663; BAND_41_4; 1.
 CC DR PROSITE; PS00664; BAND_41_5; 1.
 CC DR PROSITE; PS00665; BAND_41_6; 1.
 CC DR PROSITE; PS00666; BAND_41_7; 1.
 CC DR PROSITE; PS00667; BAND_41_8; 1.
 CC DR PROSITE; PS00668; BAND_41_9; 1.
 CC DR PROSITE; PS00669; BAND_41_10; 1.
 CC DR PROSITE; PS00670; BAND_41_11; 1.
 CC DR PROSITE; PS00671; BAND_41_12; 1.
 CC DR PROSITE; PS00672; BAND_41_13; 1.
 CC DR PROSITE; PS00673; BAND_41_14; 1.
 CC DR PROSITE; PS00674; BAND_41_15; 1.
 CC DR PROSITE; PS00675; BAND_41_16; 1.
 CC DR PROSITE; PS00676; BAND_41_17; 1.
 CC DR PROSITE; PS00677; BAND_41_18; 1.
 CC DR PROSITE; PS00678; BAND_41_19; 1.
 CC DR PROSITE; PS00679; BAND_41_20; 1.
 CC DR PROSITE; PS00680; BAND_41_21; 1.
 CC DR PROSITE; PS00681; BAND_41_22; 1.
 CC DR PROSITE; PS00682; BAND_41_23; 1.
 CC DR PROSITE; PS00683; BAND_41_24; 1.
 CC DR PROSITE; PS00684; BAND_41_25; 1.
 CC DR PROSITE; PS00685; BAND_41_26; 1.
 CC DR PROSITE; PS00686; BAND_41_27; 1.
 CC DR PROSITE; PS00687; BAND_41_28; 1.
 CC DR PROSITE; PS00688; BAND_41_29; 1.
 CC DR PROSITE; PS00689; BAND_41_30; 1.
 CC DR PROSITE; PS00690; BAND_41_31; 1.
 CC DR PROSITE; PS00691; BAND_41_32; 1.
 CC DR PROSITE; PS00692; BAND_41_33; 1.
 CC DR PROSITE; PS00693; BAND_41_34; 1.
 CC DR PROSITE; PS00694; BAND_41_35; 1.
 CC DR PROSITE; PS00695; BAND_41_36; 1.
 CC DR PROSITE; PS00696; BAND_41_37; 1.
 CC DR PROSITE; PS00697; BAND_41_38; 1.
 CC DR PROSITE; PS00698; BAND_41_39; 1.
 CC DR PROSITE; PS00699; BAND_41_40; 1.
 CC DR PROSITE; PS00700; BAND_41_41; 1.
 CC DR PROSITE; PS00701; BAND_41_42; 1.
 CC DR PROSITE; PS00702; BAND_41_43; 1.
 CC DR PROSITE; PS00703; BAND_41_44; 1.
 CC DR PROSITE; PS00704; BAND_41_45; 1.
 CC DR PROSITE; PS00705; BAND_41_46; 1.
 CC DR PROSITE; PS00706; BAND_41_47; 1.
 CC DR PROSITE; PS00707; BAND_41_48; 1.
 CC DR PROSITE; PS00708; BAND_41_49; 1.
 CC DR PROSITE; PS00709; BAND_41_50; 1.
 CC DR PROSITE; PS00710; BAND_41_51; 1.
 CC DR PROSITE; PS00711; BAND_41_52; 1.
 CC DR PROSITE; PS00712; BAND_41_53; 1.
 CC DR PROSITE; PS00713; BAND_41_54; 1.
 CC DR PROSITE; PS00714; BAND_41_55; 1.
 CC DR PROSITE; PS00715; BAND_41_56; 1.
 CC DR PROSITE; PS00716; BAND_41_57; 1.
 CC DR PROSITE; PS00717; BAND_41_58; 1.
 CC DR PROSITE; PS00718; BAND_41_59; 1.
 CC DR PROSITE; PS00719; BAND_41_60; 1.
 CC DR PROSITE; PS00720; BAND_41_61; 1.
 CC DR PROSITE; PS00721; BAND_41_62; 1.
 CC DR PROSITE; PS00722; BAND_41_63; 1.
 CC DR PROSITE; PS00723; BAND_41_64; 1.
 CC DR PROSITE; PS00724; BAND_41_65; 1.
 CC DR PROSITE; PS00725; BAND_41_66; 1.
 CC DR PROSITE; PS00726; BAND_41_67; 1.
 CC DR PROSITE; PS00727; BAND_41_68; 1.
 CC DR PROSITE; PS00728; BAND_41_69; 1.
 CC DR PROSITE; PS00729; BAND_41_70; 1.
 CC DR PROSITE; PS00730; BAND_41_71; 1.
 CC DR PROSITE; PS00731; BAND_41_72; 1.
 CC DR PROSITE; PS00732; BAND_41_73; 1.
 CC DR PROSITE; PS00733; BAND_41_74; 1.
 CC DR PROSITE; PS00734; BAND_41_75; 1.
 CC DR PROSITE; PS00735; BAND_41_76; 1.
 CC DR PROSITE; PS00736; BAND_41_77; 1.
 CC DR PROSITE; PS00737; BAND_41_78; 1.
 CC DR PROSITE; PS00738; BAND_41_79; 1.
 CC DR PROSITE; PS00739; BAND_41_80; 1.
 CC DR PROSITE; PS00740; BAND_41_81; 1.
 CC DR PROSITE; PS00741; BAND_41_82; 1.
 CC DR PROSITE; PS00742; BAND_41_83; 1.
 CC DR PROSITE; PS00743; BAND_41_84; 1.
 CC DR PROSITE; PS00744; BAND_41_85; 1.
 CC DR PROSITE; PS00745; BAND_41_86; 1.
 CC DR PROSITE; PS00746; BAND_41_87; 1.
 CC DR PROSITE; PS00747; BAND_41_88; 1.
 CC DR PROSITE; PS00748; BAND_41_89; 1.
 CC DR PROSITE; PS00749; BAND_41_90; 1.
 CC DR PROSITE; PS00750; BAND_41_91; 1.
 CC DR PROSITE; PS00751; BAND_41_92; 1.
 CC DR PROSITE; PS00752; BAND_41_93; 1.
 CC DR PROSITE; PS00753; BAND_41_94; 1.
 CC DR PROSITE; PS00754; BAND_41_95; 1.
 CC DR PROSITE; PS00755; BAND_41_96; 1.
 CC DR PROSITE; PS00756; BAND_41_97; 1.
 CC DR PROSITE; PS00757; BAND_41_98; 1.
 CC DR PROSITE; PS00758; BAND_41_99; 1.
 CC DR PROSITE; PS00759; BAND_41_100; 1.
 CC DR PROSITE; PS00760; BAND_41_101; 1.
 CC DR PROSITE; PS00761; BAND_41_102; 1.
 CC DR PROSITE; PS00762; BAND_41_103; 1.
 CC DR PROSITE; PS00763; BAND_41_104; 1.
 CC DR PROSITE; PS00764; BAND_41_105; 1.
 CC DR PROSITE; PS00765; BAND_41_106; 1.
 CC DR PROSITE; PS00766; BAND_41_107; 1.
 CC DR PROSITE; PS00767; BAND_41_108; 1.
 CC DR PROSITE; PS00768; BAND_41_109; 1.
 CC DR PROSITE; PS00769; BAND_41_110; 1.
 CC DR PROSITE; PS00770; BAND_41_111; 1.
 CC DR PROSITE; PS00771; BAND_41_112; 1.
 CC DR PROSITE; PS00772; BAND_41_113; 1.
 CC DR PROSITE; PS00773; BAND_41_114; 1.
 CC DR PROSITE; PS00774; BAND_41_115; 1.
 CC DR PROSITE; PS00775; BAND_41_116; 1.
 CC DR PROSITE; PS00776; BAND_41_117; 1.
 CC DR PROSITE; PS00777; BAND_41_118; 1.
 CC DR PROSITE; PS00778; BAND_41_119; 1.
 CC DR PROSITE; PS00779; BAND_41_120; 1.
 CC DR PROSITE; PS00780; BAND_41_121; 1.
 CC DR PROSITE; PS00781; BAND_41_122; 1.
 CC DR PROSITE; PS00782; BAND_41_123; 1.
 CC DR PROSITE; PS00783; BAND_41_124; 1.
 CC DR PROSITE; PS00784; BAND_41_125; 1.
 CC DR PROSITE; PS00785; BAND_41_126; 1.
 CC DR PROSITE; PS00786; BAND_41_127; 1.
 CC DR PROSITE; PS00787; BAND_41_128; 1.
 CC DR PROSITE; PS00788; BAND_41_129; 1.
 CC DR PROSITE; PS00789; BAND_41_130; 1.
 CC DR PROSITE; PS00790; BAND_41_131; 1.
 CC DR PROSITE; PS00791; BAND_41_132; 1.
 CC DR PROSITE; PS00792; BAND_41_133; 1.
 CC DR PROSITE; PS00793; BAND_41_134; 1.
 CC DR PROSITE; PS00794; BAND_41_135; 1.
 CC DR PROSITE; PS00795; BAND_41_136; 1.
 CC DR PROSITE; PS00796; BAND_41_137; 1.
 CC DR PROSITE; PS00797; BAND_41_138; 1.
 CC DR PROSITE; PS00798; BAND_41_139; 1.
 CC DR PROSITE; PS00799; BAND_41_140; 1.
 CC DR PROSITE; PS00800; BAND_41_141; 1.
 CC DR PROSITE; PS00801; BAND_41_142; 1.
 CC DR PROSITE; PS00802; BAND_41_143; 1.
 CC DR PROSITE; PS00803; BAND_41_144; 1.
 CC DR PROSITE; PS00804; BAND_41_145; 1.
 CC DR PROSITE; PS00805; BAND_41_146; 1.
 CC DR PROSITE; PS00806; BAND_41_147; 1.
 CC DR PROSITE; PS00807; BAND_41_148; 1.
 CC DR PROSITE; PS00808; BAND_41_149; 1.
 CC DR PROSITE; PS00809; BAND_41_150; 1.
 CC DR PROSITE; PS00810; BAND_41_151; 1.
 CC DR PROSITE; PS00811; BAND_41_152; 1.
 CC DR PROSITE; PS00812; BAND_41_153; 1.
 CC DR PROSITE; PS00813; BAND_41_154; 1.
 CC DR PROSITE; PS00814; BAND_41_155; 1.
 CC DR PROSITE; PS00815; BAND_41_156; 1.
 CC DR PROSITE; PS00816; BAND_41_157; 1.
 CC DR PROSITE; PS00817; BAND_41_158; 1.
 CC DR PROSITE; PS00818; BAND_41_159; 1.
 CC DR PROSITE; PS00819; BAND_41_160; 1.
 CC DR PROSITE; PS00820; BAND_41_161; 1.
 CC DR PROSITE; PS00821; BAND_41_162; 1.
 CC DR PROSITE; PS00822; BAND_41_163; 1.
 CC DR PROSITE; PS00823; BAND_41_164; 1.
 CC DR PROSITE; PS00824; BAND_41_165; 1.
 CC DR PROSITE; PS00825; BAND_41_166; 1.
 CC DR PROSITE; PS00826; BAND_41_167; 1.
 CC DR PROSITE; PS00827; BAND_41_168; 1.
 CC DR PROSITE; PS00828; BAND_41_169; 1.
 CC DR PROSITE; PS00829; BAND_41_170; 1.
 CC DR PROSITE; PS00830; BAND_41_171; 1.
 CC DR PROSITE; PS00831; BAND_41_172; 1.
 CC DR PROSITE; PS00832; BAND_41_173; 1.
 CC DR PROSITE; PS00833; BAND_41_174; 1.
 CC DR PROSITE; PS00834; BAND_41_175; 1.
 CC DR PROSITE; PS00835; BAND_41_176; 1.
 CC DR PROSITE; PS00836; BAND_41_177; 1.
 CC DR PROSITE; PS00837; BAND_41_178; 1.
 CC DR PROSITE; PS00838; BAND_41_179; 1.
 CC DR PROSITE; PS00839; BAND_41_180; 1.
 CC DR PROSITE; PS00840; BAND_41_181; 1.
 CC DR PROSITE; PS00841; BAND_41_182; 1.
 CC DR PROSITE; PS00842; BAND_41_183; 1.
 CC DR PROSITE; PS00843; BAND_41_184; 1.
 CC DR PROSITE; PS00844; BAND_41_185; 1.
 CC DR PROSITE; PS00845; BAND_41_186; 1.
 CC DR PROSITE; PS00846; BAND_41_187; 1.
 CC DR PROSITE; PS00847; BAND_41_188; 1.
 CC DR PROSITE; PS00848; BAND_41_189; 1.
 CC DR PROSITE; PS00849; BAND_41_190; 1.
 CC DR PROSITE; PS00850; BAND_41_191; 1.
 CC DR PROSITE; PS00851; BAND_41_192; 1.
 CC DR PROSITE; PS00852; BAND_41_193; 1.
 CC DR PROSITE; PS00853; BAND_41_194; 1.
 CC DR PROSITE; PS00854; BAND_41_195; 1.
 CC DR PROSITE; PS00855; BAND_41_196; 1.
 CC DR PROSITE; PS00856; BAND_41_197; 1.
 CC DR PROSITE; PS00857; BAND_41_198; 1.
 CC DR PROSITE; PS00858; BAND_41_199; 1.
 CC DR PROSITE; PS00859; BAND_41_200; 1.
 CC DR PROSITE; PS00860; BAND_41_201; 1.
 CC DR PROSITE; PS00861; BAND_41_202; 1.
 CC DR PROSITE; PS00862; BAND_41_203; 1.
 CC DR PROSITE; PS00863; BAND_41_204; 1.
 CC DR PROSITE; PS00864; BAND_41_205; 1.
 CC DR PROSITE; PS00865; BAND_41_206; 1.
 CC DR PROSITE; PS00866; BAND_41_207; 1.
 CC DR PROSITE; PS00867; BAND_41_208; 1.
 CC DR PROSITE; PS00868; BAND_41_209; 1.
 CC DR PROSITE; PS00869; BAND_41_210; 1.
 CC DR PROSITE; PS00870; BAND_41_211; 1.
 CC DR PROSITE; PS00871; BAND_41_212; 1.
 CC DR PROSITE; PS00872; BAND_41_213; 1.
 CC DR PROSITE; PS00873; BAND_41_214; 1.
 CC DR PROSITE; PS00874; BAND_41_215; 1.
 CC DR PROSITE; PS00875; BAND_41_216; 1.
 CC DR PROSITE; PS00876; BAND_41_217; 1.
 CC DR PROSITE; PS00877; BAND_41_218; 1.
 CC DR PROSITE; PS00878; BAND_41_219; 1.
 CC DR PROSITE; PS00879; BAND_41_220; 1.
 CC DR PROSITE; PS00880; BAND_41_221; 1.
 CC DR PROSITE; PS00881; BAND_41_222; 1.
 CC DR PROSITE; PS00882; BAND_41_223; 1.
 CC DR PROSITE; PS00883; BAND_41_224; 1.
 CC DR PROSITE; PS00884; BAND_41_225; 1.
 CC DR PROSITE; PS00885; BAND_41_226; 1.
 CC DR PROSITE; PS00886; BAND_41_227; 1.
 CC DR PROSITE; PS00887; BAND_41_228; 1.
 CC DR PROSITE; PS00888; BAND_41_229; 1.
 CC DR PROSITE; PS00889; BAND_41_230; 1.
 CC DR PROSITE; PS00890; BAND_41_231; 1.
 CC DR PROSITE; PS00891; BAND_41_232; 1.
 CC DR PROSITE; PS00892; BAND_41_233; 1.
 CC DR PROSITE; PS00893; BAND_41_234; 1.
 CC DR PROSITE; PS00894; BAND_41_235; 1.
 CC DR PROSITE; PS00895; BAND_41_236; 1.
 CC DR PROSITE; PS00896; BAND_41_237; 1.
 CC DR PROSITE; PS00897; BAND_41_238; 1.
 CC DR PROSITE; PS00898; BAND_41_239; 1.
 CC DR PROSITE; PS00899; BAND_41_240; 1.
 CC DR PROSITE; PS00900; BAND_41_241; 1.
 CC DR PROSITE; PS00901; BAND_41_242; 1.
 CC DR PROSITE; PS00902; BAND_41_243; 1.
 CC DR PROSITE; PS00903; BAND_41_244; 1.
 CC DR PROSITE; PS00904; BAND_41_245; 1.
 CC DR PROSITE; PS00905; BAND_41_246; 1.
 CC DR PROSITE; PS00906; BAND_41_247; 1.
 CC DR PROSITE; PS00907; BAND_41_248; 1.
 CC DR PROSITE; PS00908; BAND_41_249; 1.
 CC DR PROSITE; PS00909; BAND_41_250; 1.
 CC DR PROSITE; PS00910; BAND_41_251; 1.
 CC DR PROSITE; PS00911; BAND_41_252; 1.
 CC DR PROSITE; PS00912; BAND_41_253; 1.
 CC DR PROSITE; PS00913; BAND_41_254; 1.
 CC DR PROSITE; PS00914; BAND_41_255; 1.
 CC DR PROSITE; PS00915; BAND_41_256; 1.
 CC DR PROSITE; PS00916; BAND_41_257; 1.
 CC DR PROSITE; PS00917; BAND_41_258; 1.
 CC DR PROSITE; PS00918; BAND_41_259; 1.
 CC DR PROSITE; PS00919; BAND_41_260; 1.
 CC DR PROSITE; PS00920; BAND_41_261; 1.
 CC DR PROSITE; PS00921; BAND_41_262; 1.
 CC DR PROSITE; PS00922; BAND_41_263; 1.
 CC DR PROSITE; PS00923; BAND_41_264; 1.
 CC DR PROSITE; PS00924; BAND_41_265; 1.
 CC DR PROSITE; PS00925; BAND_41_266; 1.
 CC DR PROSITE; PS00926; BAND_41_267; 1.
 CC DR PROSITE; PS00927; BAND_41_268; 1.
 CC DR PROSITE; PS00928; BAND_41_269; 1.
 CC DR PROSITE; PS00929; BAND_41_270; 1.
 CC DR PROSITE; PS00930; BAND_41_271; 1.
 CC DR PROSITE; PS00931; BAND_41_272; 1.
 CC DR PROSITE; PS00932; BAND_41_273; 1.
 CC DR PROSITE; PS00933; BAND_41_274; 1.
 CC DR PROSITE; PS00934; BAND_41_275; 1.
 CC DR PROSITE; PS00935; BAND_41_276; 1.
 CC DR PROSITE; PS00936; BAND_41_277; 1.
 CC DR PROSITE; PS00937; BAND_41_278; 1.
 CC DR PROSITE; PS00938; BAND_41_279; 1.
 CC DR PROSITE; PS00939; BAND_41_280; 1.
 CC DR PROSITE; PS00940; BAND_41_281; 1.
 CC DR PROSITE; PS00941; BAND_41_282; 1.
 CC DR PROSITE; PS00942; BAND_41_283; 1.
 CC DR PROSITE; PS00943; BAND_41_284; 1.
 CC DR PROSITE; PS00944; BAND_41_285; 1.
 CC DR PROSITE; PS00945; BAND_41_286; 1.
 CC DR PROSITE; PS00946; BAND_41_287; 1.
 CC DR PROSITE; PS00947; BAND_41_288; 1.
 CC DR PROSITE; PS00948; BAND_41_289; 1.
 CC DR PROSITE; PS00949; BAND_41_290; 1.
 CC DR PROSITE; PS00950; BAND_41_291; 1.
 CC DR PROSITE; PS00951; BAND_41_292; 1.
 CC DR PROSITE; PS00952; BAND_41_293; 1.
 CC DR PROSITE; PS00953; BAND_41_294; 1.
 CC DR PROSITE; PS00954; BAND_41_295; 1.
 CC DR PROSITE; PS00955; BAND_41_296; 1.
 CC DR PROSITE; PS00956; BAND_41_297; 1.
 CC DR PROSITE; PS00957; BAND_41_298; 1.
 CC DR PROSITE; PS00958; BAND_41_299; 1.
 CC DR PROSITE; PS00959; BAND_41_300; 1.
 CC DR PROSITE; PS00960; BAND_41_301; 1.
 CC DR PROSITE; PS00961; BAND_41_302; 1.
 CC DR PROSITE; PS00962; BAND_41_303; 1.
 CC DR PROSITE; PS00963; BAND_41_304; 1.
 CC DR PROSITE; PS00964; BAND_41_305; 1.
 CC DR PROSITE; PS00965; BAND_41_306; 1.
 CC DR PROSITE; PS00966; BAND_41_307; 1.
 CC DR PROSITE; PS00967; BAND_41_308; 1.
 CC DR PROSITE; PS00968; BAND_41_309; 1.
 CC DR PROSITE; PS00969; BAND_41_310; 1.
 CC DR PROSITE; PS00970; BAND_41_311; 1.
 CC DR PROSITE; PS00971; BAND_41_312; 1.
 CC DR PROSITE; PS00972; BAND_41_313; 1.
 CC DR PROSITE; PS00973; BAND_41_314; 1.
 CC DR PROSITE; PS00974; BAND_41_315; 1.
 CC DR PROSITE; PS00975; BAND_41_316; 1.
 CC DR PROSITE; PS00976; BAND_41_317; 1.
 CC DR PROSITE; PS00977; BAND_41_318; 1.
 CC DR PROSITE; PS00978; BAND_41_319; 1.
 CC DR PROSITE; PS00979; BAND_41_320; 1.
 CC DR PROSITE; PS00980; BAND_41_321; 1.
 CC DR PROSITE; PS00981; BAND_41_322; 1.
 CC DR PROSITE; PS00982; BAND_41_323; 1.
 CC DR PROSITE; PS00983; BAND_41_324; 1.
 CC DR PROSITE; PS00984; BAND_41_325; 1.
 CC DR PROSITE; PS00985; BAND_41_326; 1.
 CC DR PROSITE; PS00986; BAND_41_327; 1.
 CC DR PROSITE; PS00987; BAND_41_328; 1.
 CC DR PROSITE; PS00988; BAND_41_329; 1.
 CC DR PROSITE; PS00989; BAND_41_330; 1.
 CC DR PROSITE; PS00990; BAND_41_331; 1.
 CC DR PROSITE; PS00991; BAND_41_332; 1.
 CC DR PROSITE; PS00992; BAND_41_333; 1.
 CC DR PROSITE; PS00993; BAND_41_334; 1.
 CC DR PROSITE; PS00994; BAND_41_335; 1.
 CC DR PROSITE; PS00995; BAND_41_336; 1.
 CC DR PROSITE; PS00996; BAND_41_337; 1.
 CC DR PROSITE; PS00997; BAND_41_338; 1.
 CC DR PROSITE; PS00998; BAND_41_339; 1.
 CC DR PROSITE; PS00999; BAND_41_340; 1.
 CC DR PROSITE; PS01000; BAND_41_341; 1.
 CC DR PROSITE; PS01001; BAND_41_342; 1.
 CC DR PROSITE; PS01002; BAND_41_343; 1.
 CC DR PROSITE; PS01003; BAND_41_344; 1.
 CC DR PROSITE; PS01004; BAND_41_345; 1.
 CC DR PROSITE; PS01005; BAND_41_346; 1.
 CC DR PROSITE; PS01006; BAND_41_347; 1.
 CC DR PROSITE; PS01007; BAND_41_348; 1.
 CC DR PROSITE; PS01008; BAND_41_349; 1.
 CC DR PROSITE; PS01009; BAND_41_350; 1.
 CC DR PROSITE; PS01010; BAND_41_351; 1.
 CC DR PROSITE; PS01011; BAND_41_352; 1.
 CC DR PROSITE; PS01012; BAND_41_353; 1.
 CC DR PROSITE; PS01013; BAND_41_354; 1.
 CC DR PROSITE; PS01014; BAND_41_355; 1.
 CC DR PROSITE; PS01015; BAND_41_356; 1.
 CC DR PROSITE; PS01016; BAND_41_357; 1.
 CC DR PROSITE; PS01017; BAND_41_358; 1.
 CC DR PROSITE; PS01018; BAND_41_359; 1.
 CC DR PROSITE; PS01019; BAND_41_360; 1.
 CC DR PROSITE; PS01020; BAND_41_361; 1.
 CC DR PROSITE; PS01021; BAND_41_362; 1.
 CC DR PROSITE; PS01022; BAND_41_363; 1.
 CC DR PROSITE; PS01023; BAND_41_364; 1.
 CC DR PROSITE; PS01024; BAND_41_365; 1.
 CC DR PROSITE; PS01025; BAND_41_366; 1.
 CC DR PROSITE; PS01026; BAND_41_367; 1.
 CC DR PROSITE; PS01027; BAND_41_368; 1.
 CC DR PROSITE; PS01028; BAND_41_369; 1.
 CC DR PROSITE; PS01029; BAND_41_370; 1.
 CC DR PROSITE; PS01030; BAND_41_371; 1.
 CC DR PROSITE; PS01031; BAND_41_372; 1.
 CC DR PROSITE; PS01032; BAND_41_373; 1.
 CC DR PROSITE; PS01033; BAND_41_374; 1.
 CC DR PROSITE; PS01034; BAND_41_375; 1.
 CC DR PROSITE; PS01035; BAND_41_376; 1.
 CC DR PROSITE; PS01036; BAND_41_377; 1.
 CC DR PROSITE; PS01037; BAND_41_378; 1.
 CC DR PROSITE; PS01038; BAND_41_379; 1.
 CC DR PROSITE; PS01039; BAND_41_380; 1.
 CC DR PROSITE; PS01040; BAND_41_381; 1.
 CC DR PROSITE; PS01041; BAND_41_382; 1.
 CC DR PROSITE; PS01042; BAND_41_383; 1.
 CC DR PROSITE; PS01043; BAND_41_384; 1.
 CC DR PROSITE; PS01044; BAND_41_385; 1.
 CC DR PROSITE; PS01045; BAND_41_386; 1.
 CC DR PROSITE; PS01046; BAND_41_387; 1.
 CC DR PROSITE; PS01047; BAND_41_388; 1.
 CC DR PROSITE; PS01048; BAND_41_389; 1.
 CC DR PROSITE; PS01049; BAND_41_390; 1.
 CC DR PROSITE; PS01050; BAND_41_391; 1.
 CC DR PROSITE; PS01051; BAND_41_392; 1.
 CC DR PROSITE; PS01052; BAND_41_393; 1.
 CC DR PROSITE; PS01053; BAND_41_394; 1.
 CC DR PROSITE; PS01054; BAND_41_395; 1.
 CC DR PROSITE; PS01055; BAND_41_396; 1.
 CC DR PROSITE; PS01056; BAND_41_

MTCTU_HELP0 STANDARD: PRT: 64 AA.

AC PS5947;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Copper-metallothionein (Cu-MT).

OS Helix pomatia (Roman snail) (Edible snail).

OC Eukaryote; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;

OC Helicaceae; Helicidae; Helix.

OX NCBI_TaxID=6536;

RN [1]

RP SEQUENCE.

RC TISSUE=Brain;

RA Dallingner R., Berger R., Hunziker P.E., Kaegi J.H.R.;

RT "Metallothionein in snail Cd and Cu metabolism."

RL Nature 388:237-238(1997).

CC -1- FUNCTION: THE METALLOTHIONEINS ARE INVOLVED IN THE CELLULAR SEQUESTRATION OF TOXIC METAL IONS AND REGULATION OF ESSENTIAL TRACE ELEMENTS. THIS ISOFORM BINDS EXCLUSIVELY COPPER.

CC -1- DOMAIN: 14 CYSTEINE RESIDUES ARE ARRANGED IN C-X-C GROUPS. THESE ARE THOUGHT TO BE THE METAL-BINDING SITES IN OTHER METALLOTHIONEINS.

CC -1- SIMILARITY: BELONGS TO THE METALLOTHIONEIN SUPERFAMILY; FAMILY 2.

CC HSP; P05106; Iuv2.

KW Metal-binding; Metal-cholate cluster; Copper; Acetylation.

DR MOD_RES 1 1

FT METAL 11 11 COPPER.

FT METAL 16 16 COPPER.

FT METAL 18 18 COPPER.

FT METAL 22 22 COPPER.

FT METAL 24 24 COPPER.

FT METAL 28 28 COPPER.

FT METAL 30 30 COPPER.

FT METAL 33 33 COPPER.

FT METAL 36 36 COPPER.

FT METAL 38 38 COPPER.

FT METAL 43 43 COPPER.

FT METAL 45 45 COPPER.

FT METAL 49 49 COPPER.

FT METAL 55 55 COPPER.

FT METAL 57 57 COPPER.

FT METAL 61 61 COPPER.

FT METAL 63 63 COPPER.

SO SEQUENCE 64 AA; 6205 MW; 96CC1998B7E12297 CRC64;

Key Match 19.2%; Score 58.5; DB 1; Length 64;

Best Local Similarity 50.0%; Pred. No. 0.96;

Matches 12; Conservative 3; Mismatches 6; Indels 3; Gaps 2;

QY 30 CTPGCKTALMGCM-KTATCHCS 52

DB 18 CGNDCKCGA--GCNDCRCSSCHS 39

RESULT 11

TL33_HUMAN

ID TL33_HUMAN STANDARD: PRT: 772 AA.

AC 004726; O9HGM5; O8WVR2;

DT 01-FEB-1994 (Rel. 28, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Transducin-like enhancer protein 3 (ESG3).

GN TL33 OR KIAA1547.

OS Homo sapiens (Human).

OC Eumetazoa; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=fetal brain;

RX MEDLINE=93265135; PubMed=1303260;

RA Stifani S., Blaumüller C.M., Redhead N.J., Hill R.E.,

RT "Human homologs of a Drosophila Enhancer of split gene product define a novel family of nuclear proteins."

RL Nat. Genet. 2:119-127(1992).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Brain;

RA MEDLINE=20450683; PubMed=10997877;

RT Nagase T., Kikuno R., Nakayama M., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."

RL DNA Res. 7:273-281(2000).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 3).

RC TISSUE=Pancreas;

RA Strausberg R.;

RL Submitted (JCT-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: NUCLEAR EFFECTOR MOLECULE.

CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; may be produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: Placenta and lung.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC -1- SIMILARITY: BELONGS TO THE GRCUCHO/TLE FAMILY OF WD-REPEAT PROTEINS.

CC EMBL; M99438; AAA61194.1; -

CC EMBL; AB046767; BAB13373.1; ALT_INT.

CC EMBL; BC015729; AAH15729.1; -

CC Genew; HGNC:11839; TLE3.

CC MIM; 600190; -

DR InterPro; IPR001680; WD40.

DR Pfam; PF00400; WD40; 6.

DR PRINTS; PR00320; GPROTEINRPT.

DR ProDom; PD000018; WD40; 1.

DR SMART; SM00320; WD40; 7.

DR PROSITE; PS00678; WD_REPEATS_1; 2.

DR PROSITE; PS50082; WD_REPEATS_2; 2.

DR PROSITE; PS50294; WD_REPEATS_REGION; 2.

KW Nuclear protein; Phosphorylation; Repeat; WD repeat;

KW Alternative splicing;

FT DOMAIN 1 131

FT DOMAIN 132 138

FT DOMAIN 199 268

FT DOMAIN 269 451

FT DOMAIN 225 228

FT DOMAIN 401 409

FT REPEAT 484 522

FT REPEAT 530 565

FT REPEAT 574 613

FT REPEAT 616 655

FT REPEAT 657 696

FT REPEAT 698 737

FT REPEAT 739 771

FT MOD_RES 240 240

FT MOD_RES 259 259

FT MOD_RES 263 263

FT MOD_RES 267 267

FT MOD_RES 342 353

FT VARSPLIC 351 353

FT VARSPLIC 417 421

FT CONFLICT 229 229

GLN-RICH.

GLY/PRO-RICH.

CCN DOMAIN.

SER/PRO-RICH.

NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

POLY-ALA.

WD 1.

WD 2.

WD 3.

WD 4.

WD 5.

WD 6.

WD 7.

PHOSPHORYLATION (BY CK2) (POTENTIAL).

PHOSPHORYLATION (BY CDC2) (POTENTIAL).

PHOSPHORYLATION (BY CDC2) (POTENTIAL).

PHOSPHORYLATION (BY CDC2) (POTENTIAL).

MISSING (IN ISOFORM 2).

MISSING (IN ISOFORM 3).

MISSING (IN ISOFORM 3).

A -> V (IN REF. 1).

```

FT CONFLICT 487 487 E -> G (IN REF. 1).
FT CONFLICT 498 498 T -> S (IN REF. 1).
FT CONFLICT 535 535 I -> M (IN REF. 1).
FT CONFLICT 541 541 L -> H (IN REF. 1).
FT CONFLICT 553 553 A -> G (IN REF. 1).
FT CONFLICT 692 692 D -> H (IN REF. 1).
FT CONFLICT 736 736 F -> S (IN REF. 1).
SQ SEQUENCE 772 AA; 83416 MW; A2A469D73BF04A43 CRC64;

Query Match 19.2%; Score 58.5; DB 1; Length 772;
Best Local Similarity 38.3%; Pred. No. 12;
Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

OY 2 STRKDFNLIVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTAT 48
Db 267 SPENGLDKARSLKRDAPTSPASVSSSTPSSKTKDL-GHNDKST 312

RESULT 12
KPGA_RAT STANDARD; PRT; 727 AA.
P51556;
DI 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Diacylglycerol kinase, alpha (EC 2.7.1.107) (Diacylglyceride kinase) (DGK-
alpha) (DAG kinase alpha) (80 kDa diacylglycerol kinase).
GN DGKA OR DAGKI OR DAGK.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=93095720; PubMed=1339302;
RT "Gene cloning, sequence, expression and in situ localization of 80
kDa diacylglycerol kinase specific to oligodendrocyte of rat brain."
RL Brain Res. Mol. Brain Res. 16:75-87(1992).
CC -1- FUNCTION: UPON CELL STIMULATION CONVERTS THE SECOND MESSENGER
DIACYLGLYCEROL INTO PHOSPHATIDATE, INITIATING THE RESYNTHESIS
OF PHOSPHATIDYLINOSITOLS AND ATTENUATING PROTEIN KINASE C
ACTIVITY.
CC -1- CATALYTIC ACTIVITY: ATP + 1,2-diacylglycerol = ADP + 1,2-
diacylglycerol 3-phosphate.
CC -1- ENZYME REGULATION: STIMULATED BY CALCIUM AND PHOSPHATIDYLSELINE.
CC -1- PHOSPHORYLATED BY PROTEIN KINASE C (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: LYMPHOCYTES AND OLIGODENDROGLIAL CELLS.
CC -1- SIMILARITY: BELONGS TO THE EUKARYOTIC DIACYLGLYCEROL KINASE
FAMILY.
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
BINDING DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; S49760; AAB24434.1; .
CC InterPro: IPR000756; DAGKA.
CC InterPro: IPR001206; DAGKC.
CC InterPro: IPR002219; DAG_PE-bind.
CC InterPro: IPR003622; DAG_Kin-cat.
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00036; ehand; 2.
CC Pfam: PF00130; DAG_PE-bind; 2.
CC Pfam: PF00609; DAGKA; 1.

```

```

DR Pfam: PF00781; DAGKC; 1.
DR PRINTS; PRO00008; DAGGEDOMAIN.
DR ProDom: PD000012; EF-hand; 1.
DR ProDom: PD002339; DAGKA; 1.
DR ProDom: PD005043; DAG_Kin-cat; 1.
DR SMART; SM00109; C1; 2.
DR SMART; SM00045; DAGKA; 1.
DR SMART; SM00046; DAGKC; 1.
DR SMART; SM00054; Efh; 2.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.
DR PROSITE; PS00018; EF_HAND; 2.
KW Transferase; kinase; Calcium-binding; Phorbol-ester binding;
KW Repeat; Multigene family.
FT CA_BIND 121 132 EF-HAND 1 (PROBABLE).
FT CA_BIND 166 177 EF-HAND 2 (PROBABLE).
FT DOMAIN 204 251 PHORBOL-ESTER AND DAG BINDING 1.
FT DOMAIN 268 317 PHORBOL-ESTER AND DAG BINDING 2.
FT DOMAIN 367 492 CATALYTIC-A (POTENTIAL).
FT DOMAIN 512 693 CATALYTIC-B (POTENTIAL).
SQ SEQUENCE 727 AA; 82198 MW; B5A248AD2F61C1D CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 727;
Best Local Similarity 24.6%; Pred. No. 15;
Matches 16; Conservative 6; Mismatches 28; Indels 15; Gaps 2;

OY 1 MSTRKDFNLIVSVSKDGSASPRITSLCTPGCKTGALMGCMKMTAT-----C 49
Db 246 MKAGPCEVSTYAKSRNDIVCPHVWVRG-----GHSGRCDRCOKIRFVHSITGLHCWC 301
OY 50 HCSIH 54
Db 302 HLEIH 306

RESULT 13
TLE3_MOUSE STANDARD; PRT; 771 AA.
ID TLE3_MOUSE
AC 008122;
DI 01-FEB-1995 (Rel. 31, Created)
DI 01-FEB-1995 (Rel. 31, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transducin-like enhancer protein 3 (ESG).
GN TLE3 OR ESG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1;
RA MEDLINE=93373944; PubMed=8365415;
RA Miyasaka H., Choudhury B.K., Hou E.W., Li S.S.-L.;
RT "Molecular cloning and expression of mouse and human cDNA encoding of
TLE3 and ESG proteins with strong similarity to Drosophila enhancer of
split groucho protein."
RL Eur. J. Biochem. 216:343-352(1993).
CC -1- FUNCTION: NUCLEAR EFFECTOR MOLECULE (BY SIMILARITY). MAY PLAY AN
IMPORTANT ROLE DURING SPERMATOGENESIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Expressed only in testis.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE GROUCHO/TLE FAMILY OF WD-REPEAT
PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC

```

DR EMBL: X73360; CA51770.1; -
DR PIR: S34162; S34162.
DR PIR: S35681; S35681.
DR MGI: 104634; Tle3.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 6.
DR PRINTS: PR00320; GPROTEINRPT.
DR PRODOM: PD000018; WD40; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 2.
DR PROSITE: PS50294; WD_REPEATS_REGION; 2.
DR Nucleic protein; Phosphorylation; Repeat; WD repeat; Spermatogenesis.
FT DOMAIN 1 130 GLN-PRO-RICH.
FT DOMAIN 131 197 GLN-PRO-RICH.
FT DOMAIN 198 267 CCN DOMAIN.
FT DOMAIN 268 450 SER/PRO-RICH.
FT DOMAIN 224 227 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT REPEAT 483 521 WD 1.
FT REPEAT 529 568 WD 2.
FT REPEAT 573 612 WD 3.
FT REPEAT 615 654 WD 4.
FT REPEAT 656 695 WD 5.
FT REPEAT 697 736 WD 6.
FT REPEAT 738 770 WD 7.
FT DOMAIN 400 408 POLY-ALA.
FT MOD_RES 239 239 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 258 258 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
FT MOD_RES 262 262 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
FT MOD_RES 266 266 PHOSPHORYLATION (BY CDC2) (POTENTIAL).
SO SEQUENCE 771 AA; 83113 MW; 42869127EAF66DD CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 771;
Best Local Similarity 38.3%; Pred. No. 16;
Matches 18; Conservative 6; Mismatches 22; Indels 1; Gaps 1;

QY 2 STRFDNLVSVSKKSGASPRTRSLCTPCKRTALMGCKMKRTAT 48
1 - - - - -
Db 266 SEPPENGLDKARGLKADAPTSPASVSSSTPSKTKDL-GHNDKST 311
1 - - - - -

RESULT 14
ENTR.MOUSE STANDARD; PRT; 1069. AA.
ID ENTR.MOUSE PRT; 1069. AA.
AC P97435;
DT 01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
Enteropitidase (EC 3.4.21.9) (Enterokinase).
PRSS7 OR ENTK.
Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Duodenum;
RX MEDLINE=98447142; PUBMED=9486188;
RA Yuan X., Zheng X., Lu D., Rubin D.C., Pung C.Y.M., Sadler J.E.;
RT "Structure of murine enterokinase (enteropeptidase) and expression in
small intestine during development.";
RT Am. J. Physiol. 274:G342-G349(1998).
CC -1- FUNCTION: RESPONSIBLE FOR INITIATING ACTIVATION OF PANCREATIC
PROTEOLYTIC PROENZYMES (TRYPsin, CHYMOTRYPsin AND CARBOXYPEPTIDASE
A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN
TURN ACTIVATES OTHER PROENZYMES INCLUDING CHYMOTRYPsinOGEN,
PROCARBOXYPEPTIDASES, AND PROELASTASES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-1-Ile-7 bond in
trypsinogen.
CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC (LIGHT) CHAIN AND A
MULTIDOMAIN (HEAVY) CHAIN LINKED BY A DISULFIDE BOND (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).

CC -1- PTM: THE CHAINS ARE DERIVED FROM A SINGLE PRECURSOR THAT IS
CLEAVED BY A TRYPSIN-LIKE PROTEASE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SEA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 MAM DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: U73378; AAB37317.1; -
CC HSSP: Q07954; 1CR8.
CC MEROPS: S01.156; -
CC MGI: 1197523; Prss7.
DR InterPro: IPR000859; CUB_domain.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR000998; MAM_domain.
DR InterPro: IPR000082; SEA_domain.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR001190; Srcr_receptot.
DR Pfam: PF00057; Idl_recept_a; 2.
DR Pfam: PF00089; trypsin; 1.
DR Pfam: PF00431; CUB; 2.
DR Pfam: PF00530; SRCR; 1.
DR Pfam: PF00629; MAM; 1.
DR Pfam: PF01390; SEA; 1.
DR PRINTS: PR00722; CHYMOTRYPsin.
DR SMART: SM00042; CUB; 2.
DR SMART: SM00192; LDLa; 2.
DR SMART: SM00137; MAM; 1.
DR SMART: SM00200; SEA; 1.
DR SMART: SM00202; SR; 1.
DR PROSITE: PS01180; CUB; 2.
DR PROSITE: PS01209; LDLRA_1; 2.
DR PROSITE: PS50068; LDLRA_2; 2.
DR PROSITE: PS00740; MAM_1; 1.
DR PROSITE: PS50060; MAM_2; 1.
DR PROSITE: PS50024; SEA; 1.
DR PROSITE: PS00420; SRCR_1; FALSE_NEG.
DR PROSITE: PS50287; SRCR_2; 1.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Signal.anchor; Glycoprotein; Hydrolase; Serine protease; Zymogen;
KW Transmembrane; Repeat.
FT CHAIN 1 829
FT CHAIN 830 1069
FT DOMAIN 1 18
FT TRANSMEM 19 47
FT DOMAIN 48 1069
FT DOMAIN 52 169
FT DOMAIN 227 268
FT DOMAIN 270 379
FT DOMAIN 387 549
FT DOMAIN 569 679
FT DOMAIN 686 724
FT DOMAIN 723 816
FT DOMAIN 830 1069
FT ACT_SITE 874 874
FT ACT_SITE 925 925
FT ACT_SITE 1021 1021
FT ACT_SITE 1021 242
FT DISULFID 236 255

CC NON-CATALYTIC CHAIN (HEAVY CHAIN).
CC CATALYTIC CHAIN (LIGHT CHAIN).
CC CYTOPLASMIC (POTENTIAL).
CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
CC EXTRACELLULAR (POTENTIAL).
CC SEA.
CC LDL-RECEPTOR CLASS A 1.
CC CUB 1.
CC MAM.
CC CUB 2.
CC LDL-RECEPTOR CLASS A 2.
CC SRCR.
CC SERINE PROTEASE.
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC BY SIMILARITY.
CC BY SIMILARITY.

FT DISULFID 249 266 BY SIMILARITY.
 FT DISULFID 688 700 BY SIMILARITY.
 FT DISULFID 695 713 BY SIMILARITY.
 FT DISULFID 707 722 BY SIMILARITY.
 FT DISULFID 817 945 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 859 875 BY SIMILARITY.
 FT DISULFID 959 1027 BY SIMILARITY.
 FT DISULFID 991 1006 BY SIMILARITY.
 FT DISULFID 1017 1045 BY SIMILARITY.
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 515 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 675 675 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 751 751 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 770 770 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 791 791 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 897 897 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 936 936 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 999 999 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1069 AA; 118735 MW; E62549E463743C3D CRC64;

Query Match 18.9%; Score 57.5; DB 1; Length 1069;
 Best Local Similarity 39.3%; Pred. No. 22;
 Matches 24; Conservative 6; Mismatches 24; Indels 7; Gaps 4;

OY 2 STKDFNIDL-VSVSKKDSGASPRITSTSL---CTPGCKTGALMGCMKMTAT-CHCSIHV 55
 DB 167 SLSDFTTAVVTISDKLTITSSPMTISASLGNLSTYVAATTSAPL-CNLSTATFATTSGHV 225
 OY 56 S 56
 DB 226 S 226

RESULT 15
 TRPF_LIPST STANDARD; PRT: 232 AA.
 ID TRPF_LIPST
 AC Q01128;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE N-(5'-phosphoribosyl)anthranilate isomerase (EC 5.3.1.24) (PRAI).
 TRP1.
 Lipomyces starkeyi.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Lipomycetaceae; Lipomyces.
 OX NCBI_TaxID=29829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCYC 1436;
 RX MEDLINE=96269934; PubMed=8662214;
 RA Bignell G.R., Bruce I.J., Evans I.H.;
 RT "Electrophoretic karyotype of the amylolytic yeast Lipomyces starkeyi
 and cloning, sequencing and chromosomal localization of its TRP1
 gene.";
 RL Curr. Genet. 30:83-88(1996).
 CC -1- CATALYTIC ACTIVITY: N-(5'-phospho-beta-D-riboseyl)-anthranilate -> 1-
 (2-carboxyphenylamino)-1-deoxy-D-ribose 5-phosphate.
 CC -1- PATHWAY: Tryptophan biosynthesis; third step.
 CC -1- SIMILARITY: BELONGS TO THE TRPF FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.

CC EMBL: Z68292; CAA92584.1;
 DR HSSP: Q56320; IDI3.
 DR InterPro: IPR001240; PRAI.
 DR Pfam: PF00697; PRAI; 1.
 KW isomerase; Tryptophan biosynthesis.
 SQ SEQUENCE 232 AA; 24625 MW; E6FF1E4EA7D7A9E0 CRC64;

Query Match 18.6%; Score 56.5; DB 1; Length 232;
 Best Local Similarity 36.7%; Pred. No. 6.2;
 Matches 11; Conservative 5; Mismatches 9; Indels 5; Gaps 1;

OY 24 ITSTSLCTPGCKTGALMGCMKMTATCHCSI 53
 DB 3 VSTSLCTPIVKI-----CGLTVEAHACAI 27

Search completed: June 7, 2003, 15:16:49
 Job time : 13 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:13:09 ; Search time 81 seconds
(without alignments)
144.996 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304
Sequence: 1 MSTKDFNLDSVSKKDSGA.....ALMGCMKATATCHCSIHYSK 57

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 671580 seqs, 206047115 residues

1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp-archaea:*
2: sp-bacteria:*
3: sp-fungi:*
4: sp-human:*
5: sp-invertebrate:*
6: sp-mammal:*
7: sp-mhc:*
8: sp-organella:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*
15: sp-virus:*
16: sp-bacteriap:*
17: sp-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	155	51.0	56	2 093GH3	093gh3 bacillus su
2	125	41.1	56	2 093GH5	093gh5 bacillus su
3	67	22.0	456	11 091V17	091v17 mus musculu
4	63	20.7	567	4 08WU13	08wu13 mus sapien
5	63	20.7	1140	4 096KG7	096kg7 homo sapien
6	62	20.4	752	10 09FKZ9	09fkz9 arabidopsis
7	61.5	20.1	536	4 08TBS6	08tbs6 homo sapien
8	61	20.1	1352	5 09V5J7	09v5j7 drosophila
9	60	19.7	456	11 0924P4	0924p4 mus musculu
10	60	19.7	773	5 08T919	08t919 drosophila
11	59.5	19.6	78	16 09PB60	09pb60 xytelia fas
12	59.5	19.6	330	5 018118	018118 caenorhabdit
13	59.5	19.6	394	5 08SMY4	08smv4 drosophila
14	59.5	19.6	804	5 09YU48	09yu48 drosophila
15	59.5	19.6	1574	11 088281	088281 rattus norv
16	59	19.4	115	4 09UG71	09ug71 homo sapien

17	58.5	19.2	63	2 08VT57	08vt57 streptococc
18	58.5	19.2	65	5 09SP49	09sp49 helix pomat
19	58.5	19.2	178	5 09UAV9	09uav9 caenorhabdit
20	58.5	19.2	365	17 097B69	097b9 thermoplasma
21	58.5	19.2	843	15 090096	090096 human immun
22	58.5	19.2	846	15 091F99	091f99 human immun
23	58.5	19.2	1546	4 09NS27	09ns27 homo sapien
24	58.5	19.2	1546	4 075445	075445 homo sapien
25	58	19.1	565	10 09SI06	09sig6 arabidopsis
26	58	19.1	613	15 08US54	08us54 human immun
27	58	19.1	648	5 09NKD7	09nk7 drosophila
28	58	19.1	689	13 08UVG4	08uv94 raja erinac
29	58	19.1	701	5 09VJU4	09vju4 drosophila
30	57.5	18.9	764	11 09JIT3	09jit3 rattus norv
31	57	18.8	57	5 09N9H2	09n9h2 venerupis d
32	57	18.8	59	5 09N9H1	09n9h1 rudilapes d
33	57	18.8	75	5 09U1N5	09u1n5 crassostrea
34	57	18.8	107	5 09NG19	09ng19 crassostrea
35	57	18.8	361	15 071036	071036 human immun
36	57	18.8	403	15 041582	041582 human immun
37	57	18.8	412	15 08Q2G0	08q2g0 human immun
38	57	18.8	656	17 08T7B5	08t7b5 methanosarc
39	56.5	18.6	262	13 09PT79	09pt79 oryzias lat
40	56.5	18.6	852	15 070010	070010 human immun
41	56	18.4	66	5 09N628	09n628 conus catus
42	56	18.4	66	5 09N625	09n625 conus catus
43	56	18.4	66	5 09NCW2	09ncw2 conus catus
44	56	18.4	506	16 08YIO5	08yio5 bruceella me
45	56	18.4	529	5 024331	024331 drosophila

ALIGNMENTS

RESULT 1	093GH3	PRELIMINARY;	PRT;	56 AA.
AC	093GH3;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Hypothetical 6.2 kDa protein.			
GN	ERISB.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;			
OC	Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=A13;			
RA	Stein T., Borchert S., Conrad B., Feesche J., Entlan K.-D.,			
RA	Hofmeister J.,			
RT	"A subtilin-like gene cluster of Bacillus subtilis A13 encodes two			
RT	(putative) lantibiotics, ericin A and ericin S."			
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AF233755; AAL15569.1; -			
KW	Hypothetical protein.			
SO	SEQUENCE 56 AA; 6241 MW; DEDDAB0892A1EBBA CRC64;			
Query Match	51.0%; Score 155; DB 2; Length 56;			
Best Local Similarity	54.7%; Pred. No. 4.5e-13;			
Matches	29; Conservative 10; Mismatches 12; Indels 2; Gaps 1;			
OY	5 DFNLDVSVSKDSGASPRITSTSLCTPGCKTGALMGCMKATATCHCSIHYSK 57			
DB	6 DFDLDVSVSKDSKITPQWKSSEVCTPGCVGLQTCFLQITTC--HISK 56			
RESULT 2	093GH5	PRELIMINARY;	PRT;	56 AA.
ID	093GH5;			
AC	093GH5;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Lantibiotic ericin Sa.
GN ERISA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A13;
RA Stein T., Borchert S., Conrad B., Feesche J., Entian K.-D.,
RT "A subtilin-like gene cluster of Bacillus subtilis A13 encodes two
RT (putative) lantibiotics, ericin A and ericin S."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF233755; AAL1567.1; -
DR InterPro: IPR001049; Gallidermin.
DR Pfam: PF02052; Gallidermin; 1.
SQ SEQUENCE 56 AA; 6195 MW; DF6F1E08B8C407D1 CRC64;

Query Match 41.3%; Score 125; DB 2; Length 56;
Best Local Similarity 60.5%; Pred. No. 4.1e-09;
Matches 23; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 5 DFNLDVSVKSDGASPRITSTLCTPGCKTGALMGC 42
DB 9 DFDLDVYKVKSDKSKIRPVLSKSLCTPGCITGPLQIC 46

RESULT 3
Q91V17 PRELIMINARY; PRT; 456 AA.
ID Q91V17
AC Q91V17;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 49.8 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BREAST TUMOR;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010331; AAH10331.1; -
DR InterPro: IPR001611; LRR.
DR Pfam: PF00560; LRR; 3.
SQ SEQUENCE 456 AA; 49816 MW; 007B782F05A357E8 CRC64;

Query Match 22.0%; Score 67; DB 11; Length 456;
Best Local Similarity 36.0%; Pred. No. 1.5;
Matches 19; Conservative 7; Mismatches 18; Indels 6; Gaps 3;

QY 8 LDIVSVKSDGASPRITSTLCTPGCKTGAL--MGCNMTATC--HCSI 53
DB 284 LSLASNLKDBGA--RLCESLLEPGQLSILWIKTSITRASCPYCSV 331

RESULT 4
Q8WUL3 PRELIMINARY; PRT; 567 AA.
ID Q8WUL3
AC Q8WUL3;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Similar to MEGF10 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-MUSCLE;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020198; AAH20198.1; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00008; EGF; 9.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00181; EGF; 11.
DR SMART; SM00180; EGF_Lam; 9.
DR PROSITE; PS00022; EGF_1; UNKNOWN_10.
DR PROSITE; PS01186; EGF_2; UNKNOWN_10.
SQ SEQUENCE 567 AA; 60797 MW; CF2FB8CDEB7CF627 CRC64;

Query Match 20.7%; Score 63; DB 4; Length 567;
Best Local Similarity 53.8%; Pred. No. 6.1;
Matches 14; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

QY 27 TSLCTPGCKTGALMGCNMTATCHCS 52
DB 149 TSRCQ--CKNGAL--CNPITGACHCA 170

RESULT 5
Q96KG7 PRELIMINARY; PRT; 1140 AA.
ID Q96KG7
AC Q96KG7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE MEGF10 protein (KIAA1780).
GN MEGF10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HIPPOCAMBUS;
RA MEDLINE=21245130; PubMed=11347906;
RX Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 8:85-95(2001).
DR EMBL; AB058676; BAB47409.1; -
DR InterPro: IPR000561; EGF-like.
DR Pfam: PF00008; EGF; 15.
DR PROSITE; PS00022; EGF_1; UNKNOWN_17.
DR PROSITE; PS01186; EGF_2; UNKNOWN_17.
SQ SEQUENCE 1140 AA; 122204 MW; 45B2FA239423895A CRC64;

Query Match 20.7%; Score 63; DB 4; Length 1140;
Best Local Similarity 53.8%; Pred. No. 12;
Matches 14; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

QY 27 TSLCTPGCKTGALMGCNMTATCHCS 52
DB 149 TSRCQ--CKNGAL--CNPITGACHCA 170

RESULT 6
Q9FKZ9 PRELIMINARY; PRT; 752 AA.
ID Q9FKZ9
AC Q9FKZ9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE Gb|AAB71479.1.
OS Arabidopsis thaliana (Mouse-ear cress).

```

```
OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asanizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,365 bp covered by twenty one
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:131-145(1998).
DR EMBL: AB010700; BAB08623.1;
SQ SEQUENCE 752 AA; 84436 MW; 1FE23D5DC461AFC2 CRC64;

Query Match
Best Local Similarity 37.7%; Score 62; DB 10; Length 752;
Matches 20; Conservative 6; Mismatches 21; Indels 6; Gaps 2;

2 STRDFNLDLV---SVSKKDSGA--SPRITSTSLCTPGCKTGALMGCNMKTAT 48
Db 652 SLKEVMTDLGAIKSKSKKDSNTNLSSQVTTTSSSTMTSEGGSSSLMMWTOT 704

RESULT 7
O8TB56 PRELIMINARY; PRT; 536 AA.
ID O8TB56
AC O8TB56;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 2610027L16 gene.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC023532; AAH25332.1;
SQ SEQUENCE 536 AA; 58248 MW; 408722C248F2851C CRC64;

Query Match
Best Local Similarity 28.3%; Score 61.5; DB 4; Length 536;
Matches 13; Conservative 9; Mismatches 23; Indels 1; Gaps 1;

11 VSVSKKDSGASPRITSTSLCTPGCKTGALMGCNMKTATCH-CSIHV 55
Db 100 LSTNRGSEMLQELGFSPLKPLCRVWALRNSLRIVACHRCGVH 145

RESULT 8
O9V5J7 PRELIMINARY; PRT; 1352 AA.
ID O9V5J7
AC O9V5J7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG12908 protein.
GN CG12908.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
```

```
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelle S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler J., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003830; AAF58809.1;
DR FlyBase; FBgn0033509; CG12908.
DR InterPro; IPR000152; ASX_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR000033; Ldl_receptor_rep.
DR InterPro; IPR003886; Nidogen_ext.
DR Pfam; PF00008; EGF; 6.
DR Pfam; PF00058; Ldl_recept_b; 3.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_like; 9.
DR SMART; SM00135; LY; 4.
DR SMART; SM00539; NTDO; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
SQ SEQUENCE 1352 AA; 149292 MW; 0C43A1DA4C0D65C CRC64;

Query Match
Best Local Similarity 38.7%; Score 61; DB 5; Length 1352;
Matches 12; Conservative 3; Mismatches 12; Indels 4; Gaps 1;

30 CTPGCKTGALMGCNMKTATCH---CSIHVS 56
Db 940 CLDYGQSDALRGCTSKPSCHLVNNGCHTAT 970

RESULT 9
O924P4 PRELIMINARY; PRT; 456 AA.
ID O924P4
AC O924P4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
```

```

DE Ribonuclease/angiogenesis inhibitor.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Melnick M.B., Comb M.J.;
RT "Mouse homolog of ribonuclease/angiogenesis inhibitor."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF071546; AAK6859.1;
DR InterPro: IPR001611; LRR.
DR Pfam: PF00560; LRR: 3.
SQ SEQUENCE 456 AA; 49626 MW; 23418247194604E5 CRC64;

Query Match 19.7%; Score 60; DB 11; Length 456;
Best Local Similarity 36.0%; Pred. No. 12;
Matches 18; Conservative 7; Mismatches 19; Indels 6; Gaps 3;

8 LDIVSVSKDGSASPRITSTSLCTPGCKTGAL--MGCNMKTATC--HCSI 53
284 LSLASLKLKDEGA--RLCESLLEPKQQLSELMIKTCSLTAAACPTFCVS 331

RESULT 10
ID 08TF919 PRELIMINARY; PRT; 773 AA.
AC 08TF919;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE AT05602P.
GN CG14982.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarini H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY075166; AAL68036.1;
DR EMBL: AY075166; AAL68036.1; 080468E1D601FCDF CRC64;
SQ SEQUENCE 773 AA; 84788 MW; 080468E1D601FCDF CRC64;

Query Match 19.7%; Score 60; DB 5; Length 773;
Best Local Similarity 34.8%; Pred. No. 21;
Matches 16; Conservative 10; Mismatches 16; Indels 4; Gaps 2;

10 LVSVSKDGSASPRITSTSLCTPGCKTGAL--MGCNMKTATC--HCSI 51
457 LVATRRDSSSTQHSANSYCGVTPAGDYSGMGSGNTTCLDRC 502

RESULT 11
ID 09PB60 PRELIMINARY; PRT; 78 AA.
AC 09PB60;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein Xf2284.
GN Xf2284.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=9A5C.
RA MEDLINE=20365717; PubMed=10910347;
RA Simpson A.D.G., Reinach F.C., Artuda P., Abreu F.A., Acencio M.,
RA Alvares M., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barro M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facchin A.P., Ferreira A.J.S., Ferreira V.C.A., Fetto J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Fromme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Martins E.M.F., Matsukuma A.Y.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sakasaka H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL: AE004040; AAF5083.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 78 AA; 8776 MW; 1CDD20E677EE32FB CRC64;

Query Match 19.6%; Score 59.5; DB 16; Length 78;
Best Local Similarity 41.2%; Pred. No. 2.4;
Matches 14; Conservative 5; Mismatches 14; Indels 1; Gaps 1;

23 RITSTSLCTPGCK--TGALMGCNMKTATCCHCSIHV 55
15 RPDVSVLTGPGCKQKQAGATLTLLRVATLYCSTHI 48

RESULT 12
ID 018118 PRELIMINARY; PRT; 330 AA.
AC 018118;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE T23FL.6 protein.
GN T23FL.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilkinson J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=99069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL: Z81129; CAB03405.1;
DR InterPro: IPR003341; DUF139.
DR Pfam: PF02363; DUF139; 7.
SQ SEQUENCE 330 AA; 36605 MW; F043B1A90D3A8FE9 CRC64;

```

```

Query Match      19.6%; Score 59.5; DB 5; Length 330;
Best Local Similarity 27.9%; Pred. No. 10;
Matches 12; Conservative 5; Mismatches 17; Indels 9; Gaps 1;

QY 23 RITSISLCTPGCKTGALMGCN-----WKATCHCHSIHVS 56
DB 275 RTTASQOQACFACSTSCNSQCNQAPAMACQPMQNSQCGQOQVVS 317

RESULT 13
OSMRY4 PRELIMINARY; PRT; 394 AA.
AC OSMRY4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE RH68811p.
GN CG17364.
OS Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;

RN RP
RC STRAIN-BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Kromm D., Farfan D., Fise E.,
RA George R., Gonzalez M., Guerin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Parras V., Park S.,
RA Patel S., Phouanenvong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Gelniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY094951; AM11304.1;
SQ SEQUENCE 394 AA; 40602 MW; 8D6ED85C06F3A3E1 CRC64;

Query Match      19.6%; Score 59.5; DB 5; Length 394;
Best Local Similarity 30.5%; Pred. No. 12;
Matches 18; Conservative 10; Mismatches 20; Indels 11; Gaps 1;

QY 10 LVSVKSKDSCGASPRITSTSLCTP-----GCKTGALMGCNKKATATCHCHSIHVS 57
DB 116 IAAVSKNSRASSRHTSGAGCSQHSRDDDDGGGGGCGTSGSCRSCESTLSMNRHLOK 174

RESULT 14
OSMRY4 PRELIMINARY; PRT; 804 AA.
AC OSMRY4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG17364 protein.
GN CG17364.
OS Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
NCBI_Taxid=7227;

RN RP
RC STRAIN-BERKELEY;
RA Adams M.D., Gelniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champagne M., Pfeiffer B.D.,
RA Wan H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beecham K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

```

```

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spter E., Spradling A.C., Stapleton M., Strong R.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AEO03536; AAF49781.1;
DR FlyBase: FBgn0036391; CG17364.
DR InterPro: IPR000217; Tubulin.
DR PROSITE: PS00227; Tubulin.
SQ SEQUENCE 804 AA; 85229 MW; BD57F6D931371345 CRC64;

Query Match      19.6%; Score 59.5; DB 5; Length 804;
Best Local Similarity 30.5%; Pred. No. 25;
Matches 18; Conservative 10; Mismatches 20; Indels 11; Gaps 1;

QY 10 LVSVKSKDSCGASPRITSTSLCTP-----GCKTGALMGCNKKATATCHCHSIHVS 57
DB 375 IAAVSKNSRASSRHTSGAGCSQHSRDDDDGGGGGCGTSGSCRSCESTLSMNRHLOK 433

RESULT 15
OSMRY4 PRELIMINARY; PRT; 1574 AA.
AC O88281;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MEGF6.
GN MEGF6.
OS Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_Taxid=10116;

RN RP
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RL like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL: AB011532; BAA32462.1;
DR HSSP: P00736; IAP0.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF-Ca.
DR Pfam: PF00008; EGF; 24.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_like; 19.
DR PROSITE: PS00010; ASX_HYDROXYL; 5.

```

Sat Jan 7 15:19:23 2003

us-10-082-618-5.rspt

Page 6

DR PROSITE, PS00022; EGF_1; UNKNOWN_23.
DR PROSITE, PS01186; EGF_2; 23.
DR PROSITE, PS01187; EGF_CA; 5.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat
SQ SEQUENCE 1574 AA; 165445 MW; 2B4853D8F77F6E7 CRC64;

Query Match	19.68;	Score 59.5;	DB 11;	length 1574;
Best Local Similarity	27.88;	Pred. No. 49;		
Matches	15;	Conservative	3;	Mismatches 9;
			Indels	27;
			Gaps	2;

```

Oy      23 RITSTSLCTPG-----CKTGLMGCKNKKTATC 51
         ||| |||
Db      757 RTGECLECPGKGTGEGDCCAGADCEGRMWGLGCQEI CPACEHGA--SCNPETGTCLC 808

```

Search completed: June 7, 2003, 15:18:31
Job time : 97 secs .

07

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:16:28 ; Search time 14 Seconds
(without alignments)
119.793 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304

Sequence: 1 MSTRDFNLIVSVSKDSGA.....ALMGCNMTATCCHSIHVK 57

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Minimum number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCtus.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	298	98.0	57	4	US-08-836-687B-20
2	295	97.0	57	4	US-08-836-687B-24
3	295	97.0	57	4	US-08-836-687B-26
4	292	96.1	57	4	US-08-836-687B-28
5	291	95.7	57	1	US-08-129-151A-2
6	291	95.7	57	2	US-08-715-579-2
7	291	95.7	57	4	US-08-836-687B-40
8	291	95.7	57	4	US-08-836-687B-42
9	289	95.1	57	4	US-08-836-687B-43
10	289	95.1	57	4	US-08-836-687B-46
11	288	94.7	57	4	US-08-836-687B-44
12	285	93.8	57	4	US-08-836-687B-45
13	285	93.8	83	3	US-08-773-731A-2
14	282.5	92.9	56	4	US-08-836-687B-41
15	280	92.1	57	4	US-08-836-687B-47
16	185	60.9	34	6	5231013-4
17	181	59.5	41	2	US-08-535-494-9
18	181	59.5	41	4	US-09-097-635-9
19	178	58.6	34	1	US-07-880-003-1
20	178	58.6	34	2	US-08-030-911-2
21	176	57.9	34	2	US-08-030-911-1
22	157	51.6	56	1	US-07-981-525-2
23	157	51.6	56	1	US-07-981-525-7
24	157	51.6	56	1	US-08-220-033-2
25	157	51.6	56	1	US-08-220-033-7
26	157	51.6	56	1	US-08-465-491-2
27	157	51.6	56	2	US-08-465-491-7

28	157	51.6	56	2	US-08-986-617-2	Sequence 2, Appl
29	157	51.6	56	2	US-08-986-617-7	Sequence 7, Appl
30	154	50.7	56	1	US-07-981-525-9	Sequence 9, Appl
31	154	50.7	56	1	US-08-220-033-9	Sequence 9, Appl
32	154	50.7	56	2	US-08-465-491-9	Sequence 9, Appl
33	154	50.7	56	2	US-08-986-617-9	Sequence 9, Appl
34	128	42.1	39	2	US-08-535-494-5	Sequence 5, Appl
35	128	42.1	39	4	US-09-097-635-5	Sequence 5, Appl
36	113	37.2	23	1	US-08-524-677-10	Sequence 10, Appl
37	113	37.2	23	2	US-08-465-491-25	Sequence 25, Appl
38	113	37.2	23	2	US-08-986-617-25	Sequence 25, Appl
39	106	34.9	34	4	US-08-836-687B-49	Sequence 49, Appl
40	105	34.5	34	4	US-08-836-687B-50	Sequence 50, Appl
41	105	34.5	34	4	US-08-836-687B-51	Sequence 51, Appl
42	104	34.2	34	2	US-08-392-625-15	Sequence 15, Appl
43	104	34.2	34	2	US-08-465-491A-15	Sequence 15, Appl
44	104	34.2	34	4	US-08-836-687B-48	Sequence 48, Appl
45	103	33.9	32	1	US-08-220-033-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-08-836-687B-20
; Sequence 20, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/770
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-20

Query Match 98.0%; Score 298; DB 4; Length 57;
Best Local Similarity 98.2%; Pred. No. 1.4e+29;
Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTRDFNLIVSVSKDSGASPRITSTSLCPGCKTGALMGCNMTATCCHSIHVK 57
DB 1 MSTRDFNLIVSVSKDSGASPRITSTSLCPGCKTGALMGCNMTATCCHSIHVK 57

RESULT 2
US-08-836-687B-24
; Sequence 24, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/770
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-24

Query Match 97.0%; Score 295; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 3.3e+29;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57
Db 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57

RESULT 3

US-08-836-687B-26
; Sequence 26, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 57
; TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-26

Query Match 97.0%; Score 295; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 3.3e-29;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57
Db 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57

RESULT 4

US-08-836-687B-28
; Sequence 28, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 57
; TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-28

Query Match 96.1%; Score 292; DB 4; Length 57;
Best Local Similarity 94.7%; Pred. No. 7.6e-29;
Matches 54; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57
Db 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57

RESULT 5

US-08-129-151A-2
; Sequence 2, Application US/08129151A
; Patent No. 5594103
; GENERAL INFORMATION:
; APPLICANT: DE VOS, Willem M.
; APPLICANT: KUIPERS, Oscar P.
; APPLICANT: KUIPERS, Oscar P.
; TITLE OF INVENTION: ANTIBIOTICS SIMILAR TO NISIN A, LACTIC
; TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR
; TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR

; TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O YOUNG & THOMPSON
; STREET: 745 South 23rd Street, Second Floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/129,151A
FILING DATE: 07-OCT-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: NL 910634
FILING DATE: 11-APR-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/NL92/00068
FILING DATE: 09-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: BO 37078

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703/521-2297
TELEFAX: 703/685-0573

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-129-151A-2
Query Match 95.7%; Score 291; DB 1; Length 57;
Best Local Similarity 96.5%; Pred. No. 1e-28;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57
Db 1 MSTKDNLDLVSVSKDGSASPRITSLCTPGCKTGALMGCMNKATATCCHSIHVK 57

RESULT 6

US-08-715-579-2
; Sequence 2, Application US/08715579
; Patent No. 5928946
; GENERAL INFORMATION:
; APPLICANT: DE VOS, Willem M.
; APPLICANT: KUIPERS, Oscar P.
; APPLICANT: KUIPERS, Oscar P.
; TITLE OF INVENTION: ANTIBIOTICS SIMILAR TO NISIN A, LACTIC
; TITLE OF INVENTION: ACID BACTERIA WHICH PRODUCE SUCH LANTIBIOTICS, METHOD FOR
; TITLE OF INVENTION: CONSTRUCTING SUCH LACTIC ACID BACTERIA AND METHOD FOR
; TITLE OF INVENTION: PRESERVING FOODSTUFFS WITH ETC.
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: C/O YOUNG & THOMPSON
; STREET: 745 South 23rd Street, Second Floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/715,579
FILING DATE: 19-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/129,151
FILING DATE: 07-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: NL 9100634
FILING DATE: 11-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00068
FILING DATE: 09-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: BO 37078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703/521-2297
TELEFAX: 703/685-0573
TELEX: 248425 EMBON
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 57 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-715-579-2

Query Match 95.7% Score 291; DB 2; Length 57;
Best Local Similarity 96.5% Pred. No. 1e-28;
Matches 55; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

RESULT 7
US-08-836-687B-40
Sequence 40, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:
APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-40

Query Match 95.7% Score 291; DB 4; Length 57;
Best Local Similarity 96.5% Pred. No. 1e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

RESULT 8
US-08-836-687B-43
Sequence 43, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:

APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-43

Query Match 95.7% Score 291; DB 4; Length 57;
Best Local Similarity 96.5% Pred. No. 1e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

RESULT 9
US-08-836-687B-42
Sequence 42, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:
APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 42
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-42

Query Match 95.1% Score 289; DB 4; Length 57;
Best Local Similarity 96.5% Pred. No. 1.8e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57
Db 1 MSTKDFNLDIVSVSKKDSGASPRITSLCTPGCKTGALMGCMNKTATCHCSIHVSK 57

RESULT 10
US-08-836-687B-46
Sequence 46, Application US/08836687B
Patent No. 6448034
GENERAL INFORMATION:
APPLICANT: Gasson, Michael John
APPLICANT: Dodd, Helen Mair
TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
FILE REFERENCE: 20747/70
CURRENT APPLICATION NUMBER: US/08/836,687B
CURRENT FILING DATE: 1995-11-20
NUMBER OF SEQ ID NOS: 51
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 57
TYPE: PRT
ORGANISM: Lactococcus sp.
US-08-836-687B-46

Query Match 95.1% Score 289; DB 4; Length 57;
Best Local Similarity 96.5% Pred. No. 1.8e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
Db 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 11
US-08-836-687B-44
; Sequence 44, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 44
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-44

Query Match 94.7%; Score 288; DB 4; Length 57;
Best Local Similarity 96.5%; Pred. No. 2.3e-28;
Matches 55; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
Db 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 12
US-08-836-687B-45
; Sequence 45, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 45
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-45

Query Match 93.8%; Score 285; DB 4; Length 57;
Best Local Similarity 94.7%; Pred. No. 5.4e-28;
Matches 54; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
Db 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 13
US-08-773-731A-2
; Sequence 2, Application US/08773731A
; Patent No. 6100056
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael J.
; APPLICANT: Dodd, Helen M.
; TITLE OF INVENTION: NISIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP

STREET: 2101 L Street N.W.
City: Washington
STATE: DC
COUNTRY: USA
ZIP: 20037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/773,731A
FILING DATE: 24-DEC-1996
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/313,123
FILING DATE: 18-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00676
FILING DATE: 01-APR-1993
APPLICATION DATA:
APPLICATION NUMBER: GB 9207267.7
FILING DATE: 02-APR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Brady, Jr., James W.
REGISTRATION NUMBER: 32,115
REFERENCE/DOCKET NUMBER: E8280.016/P016
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-785-9700
TELEFAX: 202-887-0689

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 83 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-773-731A-2

Query Match 93.8%; Score 285; DB 3; Length 83;
Best Local Similarity 94.7%; Pred. No. 8.2e-28;
Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
Db 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57

RESULT 14
US-08-836-687B-41
; Sequence 41, Application US/08836687B
; Patent No. 6448034
; GENERAL INFORMATION:
; APPLICANT: Gasson, Michael John
; APPLICANT: Dodd, Helen Mair
; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
; FILE REFERENCE: 20747/70
; CURRENT APPLICATION NUMBER: US/08/836,687B
; CURRENT FILING DATE: 1995-11-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 41
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Lactococcus sp.
US-08-836-687B-41

Query Match 92.9%; Score 282.5; DB 4; Length 56;
Best Local Similarity 96.5%; Pred. No. 1.1e-27;
Matches 55; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 57
Db 1 MSTKDFNLDVSVSKKDSGASPRITSTSLCTPGCKTGALMGCMKMTATCHCSIHVSK 56

RESULT 15

US-08-836-687B-47
 ; Sequence 47 Application US/08836687B
 ; Patent No. 6448034
 ; GENERAL INFORMATION:
 ; APPLICANT: Gasson, Michael John
 ; APPLICANT: Dodd, Helen Mair
 ; TITLE OF INVENTION: PRODUCTION OF VARIANT NISIN
 ; FILE REFERENCE: 20747/70
 ; CURRENT APPLICATION NUMBER: US/08/836,687B
 ; CURRENT FILING DATE: 1995-11-20
 ; NUMBER OF SEQ ID NOS: 51
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 47
 ; LENGTH: 57
 ; TYPE: PRT
 ; ORGANISM: Lactococcus sp.
 -08-836-687B-47

Query Match 92.1%; Score 280; DB 4; Length 57;
 Best Local Similarity 94.7%; Pred. No. 2.2e-27;
 Matches 54; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCCHSIHYSK 57
 ||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1 MSTKDFNLDLVSVSKKDSGASPRITSLCTPGCKTGALMGCMKMTATCCHSIHYSK 57
 ||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: June 7, 2003, 15:19:34
 Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2003, 15:14:21 ; Search time 20 Seconds
(without alignments)
294.235 Million cell updates/sec

Title: US-10-082-618-5

Perfect score: 304
Sequence: 1 MSTDFNLDLVSVSKDSGA.....ALMGCMKATATCHCSIHYSK 57

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

tal number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PT1_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185	60.9	34	10	US-09-030-619-206
2	185	60.9	34	10	US-09-917-340-14
3	157	51.6	56	10	US-09-030-619-208
4	73	24.0	708	9	US-10-184-644-211
5	73	24.0	708	9	US-10-184-634-211
6	70	23.0	4185	9	US-10-123-155-67
7	68.5	22.5	1570	9	US-10-184-644-335
8	68.5	22.5	1570	9	US-10-184-634-435
9	68.5	22.5	1617	9	US-10-184-644-235
10	68.5	22.5	1617	9	US-10-184-644-235
11	68.5	22.5	1875	9	US-10-123-155-317
12	68	22.4	1660	9	US-10-184-634-147
13	68	22.4	1781	9	US-10-123-155-419
14	68	22.4	2162	9	US-10-184-644-233
15	68	22.4	2162	9	US-10-184-634-233
16	68	22.4	4440	9	US-10-174-590-525
17	68	22.4	4440	9	US-10-176-758-525
18	68	22.4	4440	9	US-10-175-737-525
19	68	22.4	4440	9	US-10-175-737-525

20	68	22.4	4440	9	US-10-173-706-525	Sequence 525, App
21	68	22.4	4440	9	US-10-175-738-525	Sequence 525, App
22	68	22.4	4440	9	US-10-175-738-525	Sequence 525, App
23	68	22.4	4440	9	US-10-176-482-525	Sequence 525, App
24	68	22.4	4440	9	US-10-176-482-525	Sequence 525, App
25	68	22.4	4440	9	US-10-176-757-525	Sequence 525, App
26	68	22.4	4440	9	US-10-176-913-525	Sequence 525, App
27	68	22.4	4440	9	US-10-180-552-525	Sequence 525, App
28	68	22.4	4440	9	US-10-180-557-525	Sequence 525, App
29	68	22.4	4440	9	US-10-173-700-525	Sequence 525, App
30	68	22.4	4440	9	US-10-174-572-525	Sequence 525, App
31	68	22.4	4440	9	US-10-174-579-525	Sequence 525, App
32	68	22.4	4440	9	US-10-174-582-525	Sequence 525, App
33	68	22.4	4440	9	US-10-174-582-525	Sequence 525, App
34	68	22.4	4440	9	US-10-175-739-525	Sequence 525, App
35	68	22.4	4440	9	US-10-175-740-525	Sequence 525, App
36	68	22.4	4440	9	US-10-175-743-525	Sequence 525, App
37	68	22.4	4440	9	US-10-176-488-525	Sequence 525, App
38	68	22.4	4440	9	US-10-176-492-525	Sequence 525, App
39	68	22.4	4440	9	US-10-176-747-525	Sequence 525, App
40	68	22.4	4440	9	US-10-176-750-525	Sequence 525, App
41	68	22.4	4440	9	US-10-176-985-525	Sequence 525, App
42	68	22.4	4440	9	US-10-176-987-525	Sequence 525, App
43	68	22.4	4440	9	US-10-176-991-525	Sequence 525, App
44	68	22.4	4440	9	US-10-176-992-525	Sequence 525, App
45	68	22.4	4440	9	US-10-176-993-525	Sequence 525, App

ALIGNMENTS

```
RESULT 1
US-09-030-619-206
Sequence 206, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030.619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 206
LENGTH: 34
TYPE: PRT
ORGANISM: Lactococcus lactis
US-09-030-619-206

Query Match      60.9%  Score 185; DB 10; Length 34;
Best Local Similarity 97.1%  Pred. No. 5; Se-15;
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      24 ITSTSLCTPGCKTGALMGCMKATATCHCSIHYSK 57
Db      1 ITSTSLCTPGCKTGALMGCMKATATCHCSIHYSK 34

RESULT 2
US-09-917-340-14
Sequence 14, Application US/09917340
Patent No. US20020090369A1
GENERAL INFORMATION:
APPLICANT: Murphy, Christopher J.
APPLICANT: McNulty, Jonathan F.
APPLICANT: Reid, Ted W.
```

;; TITLE OF INVENTION: Transplant Media
;; FILE REFERENCE: TPLANT-06468
;; CURRENT APPLICATION NUMBER: US/09/917,340
;; CURRENT FILING DATE: 2001-07-29
;; PRIOR APPLICATION NUMBER: 60/221,632
;; PRIOR FILING DATE: 2000-07-28
;; PRIOR APPLICATION NUMBER: 60/249,602
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/290,932
;; PRIOR FILING DATE: 2001-05-15
;; NUMBER OF SEQ ID NOS: 96
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 14
;; LENGTH: 34
;; TYPE: PRT
;; ORGANISM: Lactococcus lactis
US-09-917-340-14

Query Match 60.9%; Score 185; DB 10; Length 34;
Best Local Similarity 97.1%; Pred. No. 5.5e-15;
Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 24 ITSTSLCTPGCKTGALMGCMNKTATCHCSIHYSK 57
DB 1 ITSISLCTPGCKTGALMGCMNKTATCHCSIHYSK 34

RESULT 3
US-09-030-619-208
; Sequence 208, Application US/09030619B
; Patent No. US20020035061A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erfile, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: MONICOL, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTI-BIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 208
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-030-619-208

Query Match 51.6%; Score 157; DB 10; Length 56;
Best Local Similarity 59.2%; Pred. No. 1.8e-11;
Matches 29; Conservative 8; Mismatches 12; Indels 0; Gaps 0;
QY 5 DFNLDIVSVSKDKSGASPRITSTSLCTPGCKTGALMGCMNKTATCHCSI 53
DB 6 DFDLDVYKYSKDKSKITTPQMKSSSLCTPGCVTGALDTCFLQTLTCKNKI 54

RESULT 4
US-10-184-644-211
; Sequence 211, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James

;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3430R1C227
;; CURRENT APPLICATION NUMBER: US/10/184,644
;; CURRENT FILING DATE: 2002-06-28
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO: 211
;; LENGTH: 708
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-184-644-211

Query Match 24.0%; Score 73; DB 9; Length 708;
Best Local Similarity 43.2%; Pred. No. 1.9;
Matches 16; Conservative 4; Mismatches 15; Indels 2; Gaps 1;
QY 18 SGASPRITSTSLCTPGCKTGALMGCMNKTAT--CHCS 52
DB 341 TGACCGCTTCTCTGCGATGACCGACCTGTGGCGCT 377

RESULT 5
US-10-184-634-211
; Sequence 211, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jilan
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO: 211
; LENGTH: 708
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-211

Query Match 24.0%; Score 73; DB 9; Length 708;
Best Local Similarity 43.2%; Pred. No. 1.9;
Matches 16; Conservative 4; Mismatches 15; Indels 2; Gaps 1;
QY 18 SGASPRITSTSLCTPGCKTGALMGCMNKTAT--CHCS 52
DB 341 TGACCGCTTCTCTGCGATGACCGACCTGTGGCGCT 377

RESULT 6
US-10-123-155-67
; Sequence 67, Application US//10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/123, 155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See Palm or File Wrapper
SEQ ID NO 67
LENGTH: 4185
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-67.

Query Match 23.0%; Score 68.5; DB 9; Length 4185;
Best Local Similarity 46.4%; Pred. No. 31;
Matches 13; Conservative 2; Mismatches 11; Indels 2; Gaps 1;

OY 25 TSTSLCTPGCKTGALMGCKNKTATCCHC 52
DB 33 TGTATCTTGCTGG--GCTATCTTCCCT 58

RESULT 7
US-10-184-644-335
Sequence 335, Application US/10184644
Publication No. US2003004930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184, 644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 335
LENGTH: 1570
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-335

Query Match 22.5%; Score 68.5; DB 9; Length 1570;
Best Local Similarity 41.2%; Pred. No. 16;
Matches 14; Conservative 2; Mismatches 15; Indels 3; Gaps 1;

OY 18 SGASPRITSTSLCTPGCKTGALMGCKNKTATCCHC 51
DB 918 TGAGCTCAATGCTGCGAAGATGCG--AGTCAC 948

RESULT 8
US-10-184-634-335
Sequence 335, Application US/10184634
Publication No. US20030068684A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184, 634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 335
LENGTH: 1570
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-634-335

Query Match 22.5%; Score 68.5; DB 9; Length 1570;
Best Local Similarity 41.2%; Pred. No. 16;
Matches 14; Conservative 2; Mismatches 15; Indels 3; Gaps 1;

OY 18 SGASPRITSTSLCTPGCKTGALMGCKNKTATCCHC 51
DB 918 TGAGCTCAATGCTGCGAAGATGCG--AGTCAC 948

RESULT 9
US-10-184-644-235
Sequence 235, Application US/10184644
Publication No. US2003004930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184, 644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 235
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-235

Query Match 22.5%; Score 68.5; DB 9; Length 1617;
Best Local Similarity 45.7%; Pred. No. 16;
Matches 16; Conservative 3; Mismatches 11; Indels 5; Gaps 2;

OY 18 SGASPRITSTSLCTPGCKTGALMGCKNKTATCCHC 52
DB 918 TGAGCTCAATGCTGCGAAGATGCG--AGTCAC 948

Db 971 AGAACAATTCCTTCCTTGA-TGC---AACACT 1000

RESULT 10

US-10-184-634-235
Sequence 235, Application US/10184634
Publication No. US20030068684A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 235
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-634-235

Query Match 22.5%; Score 68.5; DB 9; Length 1617;
Best Local Similarity 45.7%; Pred. No. 16;
Matches 16; Conservative 3; Mismatches 11; Indels 5; Gaps 2;

Qy 18 SGASPRITSTSLCTPGCKTGALMGCMKMTATC 52

Db 971 AGAACAATTCCTTCCTTGA-TGC---AACACT 1000

RESULT 11

US-10-123-155-317
Sequence 317, Application US/10123155
Publication No. US20030068794A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123,155
CURRENT FILING DATE: 2002-04-15
Prior Application removed - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 317
LENGTH: 1675
TYPE: DNA

ORGANISM: Homo Sapien
US-10-123-155-317

Query Match 22.5%; Score 68.5; DB 9; Length 1675;
Best Local Similarity 48.4%; Pred. No. 19;
Matches 15; Conservative 2; Mismatches 11; Indels 3; Gaps 1;

Qy 19 GASPRITSTSLCTPGCKTGALMGCMKMTATC 49

Db 476 GACTGCCATCACTTGCTG---GCTCCATTC 503

RESULT 12

US-10-184-644-147
Sequence 147, Application US/10184644
Publication No. US20030044930A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 147
LENGTH: 1660
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-147

Query Match 22.4%; Score 68; DB 9; Length 1660;
Best Local Similarity 37.5%; Pred. No. 19;
Matches 12; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

Qy 18 SGASPRITSTSLCTPGCKTGALMGCMKMTATC 49

Db 735 AGTGATTCATCTTGCTGCCAGCAATGAC 766

RESULT 13

US-10-184-634-147
Sequence 147, Application US/10184634
Publication No. US20030068684A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C217
CURRENT APPLICATION NUMBER: US/10/184,634
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612


```

; SEQ ID NO 147
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-634-147

```

```

Query Match
Best Local Similarity 37.5%; Score 68; DB 9; Length 1660;
Matches 12; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

```

```

OY 18 SGASPRITSTSLCTPGCKTGALMGNMKTATC 49
DB 735 ACTGATTCATCTCTGCGTGCACGCAATGAC 766

```

```

RESULT 14
US-10-123-155-419
; Sequence 419, Application US/10123155
; Publication No. US20030068794A1
GENERAL INFORMATION:

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Filvaroff, Ellen
APPLICANT: Geo, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C30
CURRENT APPLICATION NUMBER: US/10/123, 155
PRIOR APPLICATION REMOVED - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 419
LENGTH: 1781
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-155-419

```

```

Query Match
Best Local Similarity 38.2%; Score 68; DB 9; Length 1781;
Matches 13; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

```

```

OY 19 GASPRITSTSLCTPGCKTGALMGNMKTATCHCS 52
DB 163 GCATCATGACCTCGCGGACTGCTGAGACTCTCA 196

```

```

RESULT 15
US-10-184-644-233
; Sequence 233, Application US/10184644
; Publication No. US20030044930A1
GENERAL INFORMATION:

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.

```

```

; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184, 644
CURRENT FILING DATE: 2002-06-28
PRIOR APPLICATION REMOVED - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 233
LENGTH: 2162
TYPE: DNA
ORGANISM: Homo Sapien
US-10-184-644-233

```

```

Query Match
Best Local Similarity 44.1%; Score 68; DB 9; Length 2162;
Matches 15; Conservative 2; Mismatches 15; Indels 2; Gaps 1;

```

```

OY 18 SGASPRITSTSLCTPGCKTG--ALMGNMKTATC 49
DB 942 TGCCATCTATGACTTGTGACACCGCATGATC 975

```

Search completed: June 7, 2003, 15:18:51
Job time : 21 secs

